SEQUENCE LISTING

<110> Whitney, Michael
       Sanders, Pamela
       Zeh, Karin

<120> METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF SCREENING ASSAYS

<130> VSD/02-01 PROV US

<160> 64

<170> PatentIn version 3.1

<210> 1
<211> 795
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(795)
<223> ß-lactamase

<400> 1
atg agt cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag   48
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1   5   10   15

ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag   96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
                20   25   30

atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act   144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
            35   40   45

ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg   192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly
        50   55   60

caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt   240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65   70   75   80

gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta   288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85   90   95

aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc   336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
            100   105   110

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt   384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
            115   120   125

ttg cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg   432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro
        130   135   140

gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct   480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro

145				150				155				160			

```
145                     150                     155                     160
gca gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt         528
Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
            165             170             175

act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa         576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180             185             190

gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att         624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
            195             200             205

gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca         672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
            210             215             220

gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg         720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225             230             235             240

acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag         768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
            245             250             255

ata ggt gcc tca ctg att aag cat tgg                                     795
Ile Gly Ala Ser Leu Ile Lys His Trp
            260             265
```

<210> 2
<211> 265
<212> PRT
<213> Escherichia coli

<400> 2

```
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5               10              15


Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
                20              25              30


Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
                35              40              45


Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly
        50              55              60


Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65              70              75              80


Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85              90              95


Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
                100             105             110
```

```
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
    115                 120             125

Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro
    130                 135             140

Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145                 150             155                 160

Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
                165             170             175

Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180             185             190

Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
            195             200             205    .

Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
    210                 215             220

Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225                 230             235                 240

Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
                245             250             255

Ile Gly Ala Ser Leu Ile Lys His Trp
            260           .    265


<210>   3
<211>   858
<212>   DNA
<213>   Escherichia coli

<220>
<221>   CDS
<222>   (1)..(858)
<223>   ß-lactamase


<400>   3
atg aga att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca      48
Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5               10              15

ttt tgc ctt cct gtt ttt ggt cac cca gaa acg ctg gtg aaa gta aaa      96
Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys
            20              25              30

gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat     144
Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
```

```
               35                        40                        45
ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt          192
Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc          240
Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct          288
Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg          336
Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
                100                 105                 110

gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt          384
Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
            115                 120                 125

gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag          432
Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
        130                 135                 140

gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt          480
Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt          528
Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta          576
Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
                180                 185                 190

act ggc gaa cta ctt act cta gct tcc cgg caa caa tta ata gac tgg          624
Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
            195                 200                 205

atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg          672
Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct          720
Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc          768
Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat          816
Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
                260                 265                 270

aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg                  858
Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275                 280                 285
```

<210>   4

```
<211>    286
<212>    PRT
<213>    Escherichia coli

<400>    4

Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5                   10                  15


Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30


Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45


Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60


Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80


Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95


Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
                100                 105                 110


Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
            115                 120                 125


Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
        130                 135                 140


Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160


Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175


Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
                180                 185                 190


Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205


Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
210                 215                 220


Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240
```

```
Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
            245                 250                 255
```

```
Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270
```

```
Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275                 280                 285
```

```
<210>   5
<211>   795
<212>   DNA
<213>   Escherichia coli

<220>
<221>   CDS
<222>   (1)..(795)
<223>   ß-lactamase


<400>   5
atg ggg cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag      48
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5                   10                  15

ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag      96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
                20                  25                  30

atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act     144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
            35                  40                  45

ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gat gac gcc ggg     192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly
        50                  55                  60

caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt     240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65                  70                  75                  80

gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta     288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85                  90                  95

aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc     336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
            100                 105                 110

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt     384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
            115                 120                 125

ttg cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg     432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro
        130                 135                 140

gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct     480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
```

```
      145                        150                        155                              160

gta gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt        528
Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
                165                    170                   175

act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa        576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
                180                    185                   190

gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att        624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
                195                    200                   205

gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca        672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
                210                    215                   220

gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg        720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225                         230                   235                   240

acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag        768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
                245                    250                   255

ata ggt gcc tca ctg att aag cat tgg                                    795
Ile Gly Ala Ser Leu Ile Lys His Trp
                260                    265
```

<210> 6
<211> 265
<212> PRT
<213> Escherichia coli

<400> 6

```
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5                   10                  15

Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
                20                  25                  30

Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
           35                  40                  45

Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly
      50                  55                  60

Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65                  70                  75                  80

Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85                  90                  95

Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
                100                 105                 110
```

```
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
        115             120             125


Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro
    130             135             140


Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145             150             155             160


Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
            165             170             175


Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180             185             190


Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
        195             200             205


Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
    210             215             220


Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225             230             235             240


Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
            245             250             255


Ile Gly Ala Ser Leu Ile Lys His Trp
            260             265
```

```
<210>   7
<211>   792
<212>   DNA
<213>   Escherichia coli

<220>
<221>   CDS
<222>   (1)..(792)
<223>   ß-lactamase


<400>   7
atg gac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg    48
Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
1               5               10              15


ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc    96
Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
                20              25              30


ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt   144
Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe
```

```
          35                        40                        45

aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt att gac gcc ggg caa       192
Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln
    50                  55                  60

gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag       240
Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu
65                  70                  75                  80

tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga       288
Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg
                85                  90                  95

gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac       336
Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn
                100                 105                 110

tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt ttg       384
Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
        115                 120                 125

cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg gag       432
His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu
    130                 135                 140

ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct gta       480
Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val
145                 150                 155                 160

gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act       528
Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr
                165                 170                 175

cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt       576
Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val
                180                 185                 190

gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct       624
Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala
                195                 200                 205

gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca       672
Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala
    210                 215                 220

ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg       720
Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr
225                 230                 235                 240

ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag ata       768
Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile
                245                 250                 255

ggt gcc tca ctg att aag cat tgg                                       792
Gly Ala Ser Leu Ile Lys His Trp
                260
```

<210> 8
<211> 264
<212> PRT
<213> Escherichia coli

<400> 8

Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
1   5   10   15

Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
            20   25   30

Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe
        35   40   45

Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln
    50   55   60

Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu
65   70   75   80

Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg
                85   90   95

Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn
            100   105   110

Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
        115   120   125

His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu
        130   135   140

Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val
145   150   155   160

Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr
                165   170   175

Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val
            180   185   190

Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala
        195   200   205

Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala
    210   215   220

Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr
225   230   235   240

Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile
            245   250   255

```
Gly Ala Ser Leu Ile Lys His Trp
            260


<210>  9
<211>  786
<212>  DNA
<213>  Escherichia coli

<220>
<221>  CDS
<222>  (1)..(786)
<223>  ß-lactamase


<400>  9
atg aaa gat gat ttt gca aaa ctt gag gaa caa ttt gat gca aaa ctc      48
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu
1               5                   10                  15

ggg atc ttt gca ttg gat aca ggt aca aac cgg acg gta gcg tat cgg      96
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg
            20                  25                  30

ccg gat gag cgt ttt gct ttt gct tcg acg att aag gct tta act gta     144
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val
        35                  40                  45

ggc gtg ctt ttg caa cag aaa tca ata gaa gat ctg aac cag aga ata     192
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile
    50                  55                  60

aca tat aca cgt gat gat ctt gta aac tac aac ccg att acg gaa aag     240
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys
65                  70                  75                  80

cac gtt gat acg gga atg acg ctc aaa gag ctt gcg gat gct tcg ctt     288
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu
                85                  90                  95

cga tat agt gac aat gcg gca cag aat ctc att ctt aaa caa att ggc     336
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly
            100                 105                 110

gga cct gaa agt ttg aaa aag gaa ctg agg aag att ggt gat gag gtt     384
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val
        115                 120                 125

aca aat ccc gaa cga ttc gaa cca gag tta aat gaa gtg aat ccg ggt     432
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly
    130                 135                 140

gaa act cag gat acc agt aca gca aga gca ctt gtc aca agc ctt cga     480
Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg
145                 150                 155                 160

gcc ttt gct ctt gaa gat aaa ctt cca agt gaa aaa cgc gag ctt tta     528
Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu
                165                 170                 175

atc gat tgg atg aaa cga aat acc act gga gac gcc tta atc cgt gcc     576
Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala
```

```
             180                      185                         190
gga gcg gca tca tat gga acc cgg aat gac att gcc atc att tgg ccg         624
Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro
        195                     200                 205

cca aaa gga gat cct gtc ggt gtg ccg gac ggt tgg gaa gtg gct gat         672
Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp
        210                 215                 220

aaa act gtt ctt gca gta tta tcc agc agg gat aaa aag gac gcc aag         720
Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys
225                     230                 235                 240

tat gat gat aaa ctt att gca gag gca aca aag gtg gta atg aaa gcc         768
Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala
                245                 250                 255

tta aac atg aac ggc aaa                                                 786
Leu Asn Met Asn Gly Lys
                260
```

```
<210>   10
<211>   262
<212>   PRT
<213>   Escherichia coli

<400>   10

Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu
1               5                   10                  15


Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg
                20                  25                  30


Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val
            35                  40                  45


Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile
        50                  55                  60


Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys
65                  70                  75                  80


His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu
                85                  90                  95


Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly
            100                 105                 110


Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val
            115                 120                 125


Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly
        130                 135                 140
```

```
Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg
145             150             155                     160


Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu
                165                 170                 175


Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala
            180                 185                 190


Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro
        195                 200                 205


Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp
    210                 215                 220


Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys
225                 230                 235                 240


Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala
            245                 250                 255


Leu Asn Met Asn Gly Lys
            260
```

```
<210>  11
<211>  720
<212>  DNA
<213>  Aequorea victoria

<220>
<221>  CDS
<222>  (1)..(720)
<223>  mutant green fluorescent protein


<400>  11
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
                20                  25                  30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc     144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc     192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

ttc tcc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag     240
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
```

```
65                    70                    75                    80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag        288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                    90                    95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag        336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
                100                   105                   110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc        384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
                115                   120                   125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aac ctg gag tac        432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr
        130                   135                   140

aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac        480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                   150                   155                   160

ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc        528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                   170                   175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc        576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
                180                   185                   190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg        624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
                195                   200                   205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc        672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
        210                   215                   220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa        720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                   230                   235
```

```
<210>   12
<211>   239
<212>   PRT
<213>   Aequorea victoria

<400>   12

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
                20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
            35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
        50                  55                  60
```

```
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80
```

```
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                    85                  90                  95
```

```
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
                100                 105                 110
```

```
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
            115                 120                 125
```

```
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr
    130                 135                 140
```

```
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160
```

```
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175
```

```
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190
```

```
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
            195                 200                 205
```

```
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220
```

```
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

```
<210>   13
<211>   690
<212>   DNA
<213>   Anemonia majano

<220>
<221>   CDS
<222>   (1)..(690)
<223>   fluorescent protein


<400>   13
atg gct ctt tca aac aag ttt atc gga gat gac atg aaa atg acc tac      48
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
1               5                   10                  15

cat atg gat ggc tgt gtc aat ggg cat tac ttt acc gtc aaa ggt gaa      96
His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
```

```
                  20                      25                      30

ggc aac ggg aag cca tac gaa ggg acg cag act tcg act ttt aaa gtc       144
Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val
             35                      40                      45

acc atg gcc aac ggt ggg ccc ctt gca ttc tcc ttt gac ata cta tct       192
Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser
         50                      55                      60

aca gtg ttc aaa tat gga aat cga tgc ttt act gcg tat cct acc agt       240
Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser
65                      70                      75                      80

atg ccc gac tat ttc aaa caa gca ttt cct gac gga atg tca tat gaa       288
Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu
                     85                      90                      95

agg act ttt acc tat gaa gat gga gga gtt gct aca gcc agt tgg gaa       336
Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu
                100                     105                     110

ata agc ctt aaa ggc aac tgc ttt gag cac aaa tcc acg ttt cat gga       384
Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly
            115                     120                     125

gtg aac ttt cct gct gat gga cct gtg atg gcg aag aag aca act ggt       432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly
        130                     135                     140

tgg gac cca tct ttt gag aaa atg act gtc tgc gat gga ata ttg aag       480
Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys
145                     150                     155                     160

ggt gat gtc acc gcg ttc ctc atg ctg caa gga ggt ggc aat tac aga       528
Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg
                    165                     170                     175

tgc caa ttc cac act tct tac aag aca aaa aaa ccg gtg acg atg cca       576
Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro
                180                     185                     190

cca aac cat gtg gtg gaa cat cgc att gcg agg acc gac ctt gac aaa       624
Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys
            195                     200                     205

ggt ggc aac agt gtt cag ctg acg gag cac gct gtt gca cat ata acc       672
Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr
        210                     215                     220

tct gtt gtc cct ttc tga                                                690
Ser Val Val Pro Phe
225
```

```
<210>   14
<211>   229
<212>   PRT
<213>   Anemonia majano

<400>   14

Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
1                   5                   10                  15
```

His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
            20   25   30

Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val
            35   40   45

Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser
    50   55   60

Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser
65   70   75   80

Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu
                85   90   95

Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu
            100   105   110

Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly
        115   120   125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly
    130   135   140

Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys
145   150   155   160

Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg
                165   170   175

Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro
            180   185   190

Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys
        195   200   205

Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr
    210   215   220

Ser Val Val Pro Phe
225

<210> 15
<211> 696
<212> DNA
<213> Zoanthus sp

<220>

```
<221>   CDS
<222>   (1)..(696)
<223>   fluorescent protein


<400>   15
atg gct cag tca aag cac ggt cta aca aaa gaa atg aca atg aaa tac        48
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
1               5                   10                  15

cgt atg gaa ggg tgc gtc gat gga cat aaa ttt gtg atc acg gga gag        96
Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
                20                  25                  30

ggc att gga tat ccg ttc aaa ggg aaa cag gct att aat ctg tgt gtg       144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
            35                  40                  45

gtc gaa ggt gga cca ttg cca ttt gcc gaa gac ata ttg tca gct gcc       192
Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
        50                  55                  60

ttt aac tac gga aac agg gtt ttc act gaa tat cct caa gac ata gtt       240
Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
65                  70                  75                  80

gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg gac agg tct       288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser  .
                85                  90                  95

ttt ctc ttt gag gat gga gca gtt tgc ata tgt aat gca gat ata aca       336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
            100                 105                 110

gtg agt gtt gaa gaa aac tgc atg tat cat gag tcc aaa ttt tat gga       384
Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
        115                 120                 125

gtg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca gat aac       432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
    130                 135                 140

tgg gag cca tcc tgc gag aag atc ata cca gta cct aag cag ggg ata       480
Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
145                 150                 155                 160

ttg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt       528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165                 170                 175

tta cgg tgc caa ttc gac aca gtt tac aaa gca aag tct gtg cca aga       576
Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
            180                 185                 190

aag atg ccg gac tgg cac ttc atc cag cat aag ctc acc cgt gaa gac       624
Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
            195                 200                 205

cgc agc gat gct aag aat cag aaa tgg cat ctg aca gaa cat gct att       672
Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
    210                 215                 220

gca tcc gga tct gca ttg ccc tga                                       696
Ala Ser Gly Ser Ala Leu Pro
```

225   230

<210>   16
<211>   231
<212>   PRT
<213>   Zoanthus sp

<400>   16

Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
1   5   10   15

Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
                20   25   30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
            35   40   45

Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
        50   55   60

Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
65   70   75   80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser
                85   90   95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
            100   105   110

Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
        115   120   125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
        130   135   140

Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
145   150   155   160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165   170   175

Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
            180   185   190

Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
            195   200   205

Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
        210   215   220

Ala Ser Gly Ser Ala Leu Pro
225   230

<210>   17
<211>   696
<212>   DNA
<213>   Zoanthus sp

<220>
<221>   CDS
<222>   (1)..(696)
<223>   fluorescent protein

<400>   17
atg gct cat tca aag cac ggt cta aaa gaa gaa atg aca atg aaa tac   48
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr
1   5   10   15

cac atg gaa ggg tgc gtc aac gga cat aaa ttt gtg atc acg ggc gaa   96
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu
                20   25   30

ggc att gga tat ccg ttc aaa ggg aaa cag act att aat ctg tgt gtg   144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val
            35   40   45

atc gaa ggg gga cca ttg cca ttt tcc gaa gac ata ttg tca gct ggc   192
Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly
        50   55   60

ttt aag tac gga gac agg att ttc act gaa tat cct caa gac ata gta   240
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val
65   70   75   80

gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg ggc agg tct   288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser
                85   90   95

ttt ctc ttt gag gat gga gca gtc tgc ata tgc aat gta gat ata aca   336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr
                100   105   110

gtg agt gtc aaa gaa aac tgc att tat cat aag agc ata ttt aat gga   384
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly
            115   120   125

atg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca act aac   432
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn
            130   135   140

tgg gaa gca tcc tgc gag aag atc atg cca gta cct aag cag ggg ata   480
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile
145   150   155   160

ctg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt   528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165   170   175

tac cgg tgc cag ttc gac aca gtt tac aaa gca aag tct gtg cca agt   576
Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser

```
                180                      185                        190
aag atg ccg gag tgg cac ttc atc cag cat aag ctc ctc cgt gaa gac          624
Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp
        195                 200                 205

cgc agc gat gct aag aat cag aag tgg cag ctg aca gag cat gct att          672
Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile
    210                 215                 220

gca ttc cct tct gcc ttg gcc tga                                          696
Ala Phe Pro Ser Ala Leu Ala
225                 230
```

<210> 18
<211> 231
<212> PRT
<213> Zoanthus sp

<400> 18

```
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr
1               5                   10                  15

His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu
            20                  25                  30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val
        35                  40                  45

Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly
    50                  55                  60

Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val
65                  70                  75                  80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser
                85                  90                  95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr
            100                 105                 110

Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly
            115                 120                 125

Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn
        130                 135                 140

Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile
145                 150                 155                 160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165                 170                 175
```

Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser
       180   185   190

Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp
       195   200   205

Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile
   210   215   220

Ala Phe Pro Ser Ala Leu Ala
225   230

<210> 19
<211> 699
<212> DNA
<213> Discosoma striata

<220>
<221> CDS
<222> (1)..(699)
<223> fluorescent protein

<400> 19
atg agt tgt tcc aag agt gtg atc aag gaa gaa atg ttg atc gat ctt   48
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu
1   5   10   15

cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa   96
His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
           20   25   30

gga aaa gga cag cct aat gaa ggc acc aat acc gtc acg ctc gag gtt   144
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val
           35   40   45

acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa   192
Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
   50   55   60

ttt cag tat gga aac aag gca ttt gtc cac cac cct gac aac ata cat   240
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His
65   70   75   80

gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cgg tcc   288
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
              85   90   95

atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt   336
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
           100   105   110

ttg aca ggc aac tgt ttc tac tac gac atc aag ttc act ggc ttg aac   384
Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn
           115   120   125

ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa   432
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu

```
            130                    135                    140
ccg agc act gag cgt ttg tat cct cgt gat ggt gtg ttg ata gga gac        480
Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
145                 150                 155                 160

atc cat cat gct ctg aca gtt gaa gga ggt ggt cat tac gca tgt gac        528
Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp
                    165                 170                 175

att aaa act gtt tac agg gcc aag aag gcc gcc ttg aag atg cca ggg        576
Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly
                180                 185                 190

tat cac tat gtt gac acc aaa ctg gtt ata tgg aac aac gac aaa gaa        624
Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu
                195                 200                 205

ttc atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cac cat ccg        672
Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro
            210                 215                 220

ttc tat gag cca aag aag gat aag taa                                    699
Phe Tyr Glu Pro Lys Lys Asp Lys
225                 230
```

```
<210>   20
<211>   232
<212>   PRT
<213>   Discosoma striata

<400>   20
```

```
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu
1               5                   10                  15


His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
            20                  25                  30


Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val
        35                  40                  45


Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
    50                  55                  60


Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His
65                  70                  75                  80


Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
                85                  90                  95


Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
            100                 105                 110


Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn
            115                 120                 125
```

```
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu
    130             135             140


Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
145             150             155             160


Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp
                165             170             175


Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly
                180             185             190


Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu
        195             200             205


Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro
    210             215             220


Phe Tyr Glu Pro Lys Lys Asp Lys
225             230
```

```
<210>   21
<211>   678
<212>   DNA
<213>   Discosoma sp "red"

<220>
<221>   CDS
<222>   (1)..(678)
<223>   fluorescent protein


<400>   21
atg agg tct tcc aag aat gtt atc aag gag ttc atg agg ttt aag gtt      48
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1               5                   10                  15

cgc atg gaa gga acg gtc aat ggg cac gag ttt gaa ata gaa ggc gaa      96
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
                20                  25                  30

gga gag ggg agg cca tac gaa ggc cac aat acc gta aag ctt aag gta     144
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
        35                  40                  45

acc aag ggg gga cct ttg cca ttt gct tgg gat att ttg tca cca caa     192
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
    50                  55                  60

ttt cag tat gga agc aag gta tat gtc aag cac cct gcc gac ata cca     240
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65                  70                  75                  80

gac tat aaa aag ctg tca ttt cct gaa gga ttt aaa tgg gaa agg gtc     288
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
```

```
                   85                        90                        95
atg aac ttt gaa gac ggt ggc gtc gtt act gta acc cag gat tcc agt        336
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
            100                 105                 110

ttg cag gat ggc tgt ttc atc tac aag gtc aag ttc att ggc gtg aac        384
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
            115                 120                 125

ttt cct tcc gat gga cct gtt atg caa aag aag aca atg ggc tgg gaa        432
Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
            130                 135                 140

gcc agc act gag cgt ttg tat cct cgt gat ggc gtg ttg aaa gga gag        480
Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145                 150                 155                 160

att cat aag gct ctg aag ctg aaa gac ggt ggt cat tac cta gtt gaa        528
Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
                165                 170                 175

ttc aaa agt att tac atg gca aag aag cct gtg cag cta cca ggg tac        576
Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
                180                 185                 190

tac tat gtt gac tcc aaa ctg gat ata aca agc cac aac gaa gac tat        624
Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
                195                 200                 205

aca atc gtt gag cag tat gaa aga acc gag gga cgc cac cat ctg ttc        672
Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
    210                 215                 220

ctt taa                                                                 678
Leu
225


<210>   22
<211>   225
<212>   PRT
<213>   Discosoma sp "red"

<400>   22

Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1               5                   10                  15


Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
                20                  25                  30


Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
            35                  40                  45


Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
    50                  55                  60


Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65                  70                  75                  80
```

```
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
                85                90                95


Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
            100                105                110


Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
            115                120                125


Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
        130                135                140


Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145                150                155                160


Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
                165                170                175


Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
            180                185                190


Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
            195                200                205


Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
    210                215                220


Leu
225
```

```
<210>   23
<211>   801
<212>   DNA
<213>   Clavularia sp

<220>
<221>   CDS
<222>   (1)..(801)
<223>   fluorescent protein


<400>   23
atg aag tgt aaa ttt gtg ttc tgc ctg tcc ttc ttg gtc ctc gcc atc      48
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1               5                   10                  15

aca aac gcg aac att ttt ttg aga aac gag gct gac tta gaa gag aag      96
Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
                20                  25                  30

aca ttg aga ata cca aaa gct cta acc acc atg ggt gtg att aaa cca     144
Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
```

```
              35                      40                      45

gac atg aag att aag ctg aag atg gaa gga aat gta aac ggg cat gct         192
Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
    50                  55                  60


ttt gtg atc gaa gga gaa gga gaa gga aag cct tac gat ggg aca cac         240
Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
65                  70                  75                  80


act tta aac ctg gaa gtg aag gaa ggt gcg cct ctg cct ttt tct tac        288
Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr
                85                  90                  95


gat atc ttg tca aac gcg ttc cag tac gga aac aga gca ttg aca aaa        336
Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys
                100                 105                 110


tac cca gac gat ata gca gac tat ttc aag cag tcg ttt ccc gag gga        384
Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly
        115                 120                 125


tat tcc tgg gaa aga acc atg act ttt gaa gac aaa ggc att gtc aaa        432
Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys
    130                 135                 140


gtg aaa agt gac ata agc atg gag gaa gac tcc ttt atc tat gaa att        480
Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile
145                 150                 155                 160


cgt ttt gat ggg atg aac ttt cct ccc aat ggt ccg gtt atg cag aaa        528
Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys
                165                 170                 175


aaa act ttg aag tgg gaa cca tcc act gag att atg tac gtg cgt gat        576
Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp
                180                 185                 190


gga gtg ctg gtc gga gat att agc cat tct ctg ttg ctg gag gga ggt        624
Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly
                195                 200                 205


ggc cat tac cga tgt gac ttc aaa agt att tac aaa gca aaa aaa gtt        672
Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val
    210                 215                 220


gtc aaa ttg cca gac tat cac ttt gtg gac cat cgc att gag atc ttg        720
Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu
225                 230                 235                 240


aac cat gac aag gat tac aac aaa gta acg ctg tat gag aat gca gtt        768
Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val
                245                 250                 255


gct cgc tat tct ttg ctg cca agt cag gcc tag                            801
Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala
                260                 265
```

<210>   24
<211>   266
<212>   PRT
<213>   Clavularia sp

<400> 24

Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1   5   10   15

Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
            20   25   30

Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
        35   40   45

Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
    50   55   60

Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
65   .   70   75   80

Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr
                85   90   95

Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys
            100   105   110

Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly
        115   120   125

Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys
    130   135   140

Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile
145   150   155   160

Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys
            165   170   175

Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp
            180   185   190

Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly
        195   200   205

Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val
    210   215   220

Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu
225   230   235   240

Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val
                245   250   255

Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala
            260   265

<210>   25
<211>   881
<212>   PRT
<213>   Saccharomyces cerevisiae

<400>   25

Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu
1   5   10   15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
            20   25   30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
            35   40   45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
    50   55   60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
65   70   75   80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
                85   90   95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
                100   105   110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
            115   120   125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu
    130   135   140

Thr Val Ser Ile Asp Ser Ala Ala His His Asp Asn Ser Thr Ile Pro
145   150   155   160

Leu Asp Phe Met Pro Arg Asp Ala Leu His Gly Phe Asp Trp Ser Glu
                165   170   175

Glu Asp Asp Met Ser Asp Gly Leu Pro Phe Leu Lys Thr Asp Pro Asn
                180   185   190

Asn Asn Gly Phe Phe Gly Asp Gly Ser Leu Leu Cys Ile Leu Arg Ser
            195   200   205

Ile Gly Phe Lys Pro Glu Asn Tyr Thr Asn Ser Asn Val Asn Arg Leu
    210   215   220

Pro Thr Met Ile Thr Asp Arg Tyr Thr Leu Ala Ser Arg Ser Thr Thr
225   230   235   240

Ser Arg Leu Leu Gln Ser Tyr Leu Asn Asn Phe His Pro Tyr Cys Pro
                245   250   255

Ile Val His Ser Pro Thr Leu Met Met Leu Tyr Asn Asn Gln Ile Glu
                260   265   270

Ile Ala Ser Lys Asp Gln Trp Gln Ile Leu Phe Asn Cys Ile Leu Ala
            275   280   285

Ile Gly Ala Trp Cys Ile Glu Gly Glu Ser Thr Asp Ile Asp Val Phe
    290   295   300

Tyr Tyr Gln Asn Ala Lys Ser His Leu Thr Ser Lys Val Phe Glu Ser
305   310   315   320

Gly Ser Ile Ile Leu Val Thr Ala Leu His Leu Leu Ser Arg Tyr Thr
            325   330   335

Gln Trp Arg Gln Lys Thr Asn Thr Ser Tyr Asn Phe His Ser Phe Ser
            340   345   350

Ile Arg Met Ala Ile Ser Leu Gly Leu Asn Arg Asp Leu Pro Ser Ser
            355   360   365

Phe Ser Asp Ser Ser Ile Leu Glu Gln Arg Arg Arg Ile Trp Trp Ser
    370   375   380

Val Tyr Ser Trp Glu Ile Gln Leu Ser Leu Leu Tyr Gly Arg Ser Ile
385   390   395   400

Gln Leu Ser Gln Asn Thr Ile Ser Phe Pro Ser Ser Val Asp Asp Val
                405   410   415

Gln Arg Thr Thr Thr Gly Pro Thr Ile Tyr His Gly Ile Ile Glu Thr
            420   425   430

Ala Arg Leu Leu Gln Val Phe Thr Lys Ile Tyr Glu Leu Asp Lys Thr
            435   440   445

Val Thr Ala Glu Lys Ser Pro Ile Cys Ala Lys Lys Cys Leu Met Ile
450   455   460

Cys Asn Glu Ile Glu Glu Val Ser Arg Gln Ala Pro Lys Phe Leu Gln
465   470   475   480

Met Asp Ile Ser Thr Thr Ala Leu Thr Asn Leu Leu Lys Glu His Pro
                485   490   495

Trp Leu Ser Phe Thr Arg Phe Glu Leu Lys Trp Lys Gln Leu Ser Leu
            500   505   510

Ile Ile Tyr Val Leu Arg Asp Phe Phe Thr Asn Phe Thr Gln Lys Lys
        515   520   525

Ser Gln Leu Glu Gln Asp Gln Asn Asp His Gln Ser Tyr Glu Val Lys
    530   535   540

Arg Cys Ser Ile Met Leu Ser Asp Ala Ala Gln Arg Thr Val Met Ser
545   550   555   560

Val Ser Ser Tyr Met Asp Asn His Asn Val Thr Pro Tyr Phe Ala Trp
            565   570   575

Asn Cys Ser Tyr Tyr Leu Phe Asn Ala Val Leu Val Pro Ile Lys Thr
            580   585   590

Leu Leu Ser Asn Ser Lys Ser Asn Ala Glu Asn Asn Glu Thr Ala Gln
    595   600   605

Leu Leu Gln Gln Ile Asn Thr Val Leu Met Leu Leu Lys Lys Leu Ala
    610   615   620

Thr Phe Lys Ile Gln Thr Cys Glu Lys Tyr Ile Gln Val Leu Glu Glu
625   630   635   640

Val Cys Ala Pro Phe Leu Leu Ser Gln Cys Ala Ile Pro Leu Pro His
            645   650   655

Ile Ser Tyr Asn Asn Ser Asn Gly Ser Ala Ile Lys Asn Ile Val Gly
            660   665   670

Ser Ala Thr Ile Ala Gln Tyr Pro Thr Leu Pro Glu Glu Asn Val Asn
            675   680   685

Asn Ile Ser Val Lys Tyr Val Ser Pro Gly Ser Val Gly Pro Ser Pro
    690   695   700

Val Pro Leu Lys Ser Gly Ala Ser Phe Ser Asp Leu Val Lys Leu Leu
705   710   715   720

```
Ser Asn Arg Pro Pro Ser Arg Asn Ser Pro Val Thr Ile Pro Arg Ser
            725             730             735

Thr Pro Ser His Arg Ser Val Thr Pro Phe Leu Gly Gln Gln Gln Gln
            740             745             750

Leu Gln Ser Leu Val Pro Leu Thr Pro Ser Ala Leu Phe Gly Gly Ala
            755             760             765

Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe Thr
            770             775             780

Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr Asn
785             790             795             800

Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp Asn
            805             810             815

Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn Asn
            820             825             830

Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn Ala
            835             840             845

Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp Val
    850             855             860

Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys Lys
865             870             875             880

Glu
```

```
<210>   26
<211>   202
<212>   PRT
<213>   Saccharomyces cerevisiae

<400>   26

Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg
1               5               10              15

Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala
                20              25              30

Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys
            35              40              45
```

Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg
   50   55   60

Gly Ile Arg Leu Leu Gln Glu Glu Glu Gly Leu Pro Leu Val Gly
65   70   75   80

Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly
          85   90   95

His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu
        100   105   110

Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly
      115   120   125

Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val
   130   135   140

Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys
145   150   155   160

Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro
        165   170   175

Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala
        180   185   190

Val Gly Val Ile Arg Asn Gly Asp Trp Leu
      195   200

<210> 27
<211> 17
<212> DNA
<213> unidentified

<220>
<221> misc_feature
<222> (4)..(4)
<223> any purine

<220>
<221> misc_feature
<222> (1)..(17)
<223> DNA sequence that is bound by GAL4

<220>
<221> misc_feature
<222> (5)..(5)
<223> any nucleotide

<220>

```
<221>  misc_feature
<222>  (6)..(6)
<223>  any nucleotide


<220>
<221>  misc_feature
<222>  (7)..(7)
<223>  any purine


<220>
<221>  misc_feature
<222>  (9)..(9)
<223>  any pyrimidine


<220>
<221>  misc_feature
<222>  (11)..(11)
<223>  any pyrimidine


<220>
<221>  misc_feature
<222>  (10)..(10)
<223>  any nucleotide


<220>
<221>  misc_feature
<222>  (12)..(12)
<223>  any nucleotide


<220>
<221>  misc_feature
<222>  (14)..(14)
<223>  any nucleotide


<400>  27
cggnnnncnn nncnccg                                              17


<210>  28
<211>  14
<212>  DNA
<213>  unidentified

<220>
<221>  misc_feature
<222>  (6)..(9)
<223>  any nucleotide


<400>  28
cgaacnnnng ttcg                                                 14


<210>  29
<211>  25
<212>  DNA
<213>  Artificial
```

```
<220>
<223>  primer

<400>  29
ctagacgttg acattgatta ttgac                                    25


<210>  30
<211>  25
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  30
tctagagcca gtaagcagtg ggttc                                    25


<210>  31
<211>  28
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  31
gcggccgccc attgcattgg gattggtc                                 28


<210>  32
<211>  26
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  32
gcggccgccc tccaagtctt tatctg                                   26


<210>  33
<211>  28
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  33
ggtaccggct cgtagagaaa tatgaacc                                 28


<210>  34
<211>  26
<212>  DNA
<213>  Artificial

<220>
<223>  primer
```

```
<400>  34
ctcgagagag actgaatttc cctttt                                          26


<210>  35
<211>  30
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  35
gcggccgcga aggaagatag ccaaggctta                                      30


<210>  36
<211>  32
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  36
gcggccgctt aaagatgcca cactgctggt ct                                   32


<210>  37
<211>  26
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  37
gtcgacacct gtcggctact gggata                                          26


<210>  38
<211>  26
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  38
gatatcgtgg gactctgttt ccagca                                          26


<210>  39
<211>  6795
<212>  DNA
<213>  Artificial

<220>
<223>  recombinant DNA vector

<220>
<221>  misc_feature
<222>  (892)..(892)
<223>  unknown nucleotide
```

<400> 39

```
aagtccggca cctcgtgcac gcggatttcg gctccaacaa tgtcctgacg gacaatggcc      60

gcataacagc ggtcattgac tggagcgagg cgatgttcgg ggattcccaa tacgaggtcg     120

ccaacatctt cttctggagg ccgtggttgg cttgtatgga gcagcagacg cgctacttcg     180

agcggaggca tccggagctt gcaggatcgc cgcggctccg ggcgtatatg ctccgcattg     240

gtcttgacca actctatcag agcttggttg acggcaattt cgatgatgca gcttgggcgc     300

agggtcgatg cgacgcaatc gtccgatccg gagccgggac tgtcgggcgt acacaaatcg     360

cccgcagaag cgcggccgtc tggaccgatg gctgtgtaga agtactcgcc gatagtggaa     420

accgacgccc cagcactcgt ccgagggcaa aggaatgcct gagaaaggaa gtgagctgta     480

aaggctgagc tctctctctg acgtatgtag cctctggtta gcttcgtcac tcactgttct     540

tgactcagca tggcaatctg atgaaatccc agctgtaagt ctgcataaat tgatgatcta     600

ttaaacaata aagatgtcca ctaaaatgga agtttttac tgtcatactt tgttaagaag     660

ggtgagaaca gagtacctac attttgaatg gaaggattgg agctacgggg gtgggggtgg    . 720

gggtgggatt agataaatgc ctgctcttta ctgaaggctc tttactattg ctttatgata     780

atgtttcata gttggatatc ataatttaaa caagcaaaac caaattaagg gccagctcat     840

tcctccactc atgatctata gatctataga tctctcgtgg gatcattgtt tntctgatcc     900

actggaagct tatcgatacc gtcgacctcg agggggggcc cggtacccag cttttgttcc     960

ctttagtgag ggttaattgc gcgcttggcg taatcatggt catagctgtt tcctgtgtga    1020

aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc    1080

tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc    1140

cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc    1200

ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt    1260

cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca    1320

ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa    1380

aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat    1440

cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc    1500

cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc    1560

gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt    1620

tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac    1680

cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg    1740

ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca    1800

gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc    1860
```

```
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa    1920

accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg cagaaaaaaa    1980

ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac    2040

tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttttta    2100

aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt    2160

taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata    2220

gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc    2280

agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac    2340

cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag    2400

tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac    2460

gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc    2520

agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg    2580

gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc    2640

atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgctttttct    2700

gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc    2760

tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc    2820

atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc    2880

agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc    2940

gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca    3000

cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt    3060

tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt    3120

ccgcgcacat ttccccgaaa agtgccacct aaattgtaag cgttaatatt ttgttaaaat    3180

tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa    3240

tcccttataa atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttggaaca    3300

agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg    3360

gcgatggccc actacgtgaa ccatcaccct aatcaagttt tttggggtcg aggtgccgta    3420

aagcactaaa tcggaaccct aaagggagcc cccgatttag agcttgacgg ggaaagccgg    3480

cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg cgctggcaa    3540

gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc gcttaatgcg ccgctacagg    3600

gcgcgtccca ttcgccattc aggctgcgca actgttggga agggcgatcg gtgcgggcct    3660

cttcgctatt acgccagctg gcgaaggggg gatgtgctgc aaggcgatta agttgggtaa    3720

cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgagcgc gcgtaatacg    3780
```

```
actcactata gggcgaattg gagctccacc gcggtgcggc cgggccatgc aggccacgac    3840

atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc    3900

tttatttgtg aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa    3960

caagttcttg tacagctcgt ccatgccgag agtgatcccg gcggcggtca cgaactccag    4020

caggaccatg tgatcgcgct tctcgttggg gtctttgctc agggcggact ggtagctcag    4080

gtagtggttg tcgggcagca gcacggggcc gtcgccgatg ggggtgttct gctggtagtg    4140

gtcggcgagc tgcacgctgc cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat    4200

gccgttcttc tgcttgtcgg ccatgatata gacgttgtgg ctgttgtagt tgtactccag    4260

cttgtgcccc aggatgttgc cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt    4320

caccagggtg tcgccctcga acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt    4380

gaagaagatg gtgcgctcct ggacgtagcc ttcgggcatg cggacttga agaagtcgtg    4440

ctgcttcatg tggtcggggt agcgggcgaa gcactgcagg ccgtagccga aggtggtcac    4500

gagggtgggc cagggcacgg gcagcttgcc ggtggtgcag atgaacttca gggtcagctt    4560

gccgtaggtg gcatcgccct cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc    4620

gccgtccagc tcgaccagga tgggcaccac cccggtgaac agctcctcgc ccttgctcac    4680

catggtggcg ttagcttgat ttgacagtgg ctgggggttg cgccgccggg ttttatagga    4740

agccacagcg gccactcgag ccataaaagg caactttagg aacggcgggg ggtgattgga    4800

ttcgagtcgt ttattcaccg gccttgccgc acagtgcagc attttttac cccctctccc    4860

ctcctttgc gggggaaaaa aaaaaaaaa aaaaaaaag gagaagagaa aaaagcgag    4920

cgagagagaa agcgagattg aggaagagga tgaagagttt ggcgatgggt gctggttccg    4980

taggcccaga tggacaagaa tagcccccgc ccttgcggac agtatcccat tcagtgactc    5040

agatcagatc aagcggccgc tctagaacta gtggatccgg ataagccagt aagcagtggg    5100

ttctctagtt agccagagag ctctgcttat atagacctcc caccgtacac gcctaccgcc    5160

catttgcgtc aatggggcgg agttgttacg acattttgga aagtcccgtt gattttggtg    5220

ccaaaacaaa ctcccattga cgtcaatggg gtggagactt ggaaatcccc gtgagtcaaa    5280

ccgctatcca cgcccattga tgtactgcca aaaccgcatc accatggtaa tagcgatgac    5340

taatacgtag atgtactgcc aagtaggaaa gtcccataag gtcatgtact gggcataatg    5400

ccaggcgggc catttaccgt cattgacgtc aataggggc gtacttggca tatgatacac    5460

ttgatgtact gccaagtggg cagtttaccg taaatactcc acccattgac gtcaatggaa    5520

agtccctatt ggcgttacta tgggaacata cgtcattatt gacgtcaatg gcggggggtc    5580

gttgggcggt cagccaggcg ggccatttac cgtaagttat gtaacgcgga actccatata    5640

tgggctatga actaatgacc ccgtaattga ttactattaa taactagtca ataatcaatg    5700
```

```
tcaaccccgg gctgcaggaa ttctaccggg tagggggaggc gcttttccca aggcagtctg    5760

gagcatgcgc tttagcagcc ccgctggcac ttggcgcatc acaagtggcc tctggcctcg    5820

cacacattcc acatccaccg gtagcgccaa ccggctccct tctttggtgg ccccttcgcg    5880

ccaccttcta ctcctcccct agtcaggaag ttcccccccg ccccgcagct cgcgtcgtgc    5940

aggacgtgac aaatggaagt agcacgtctc actagtctcg tgcagatgga caagcaccgc    6000

tgagcaatgg aagcgggtag gcctttgggg cagcggccaa tagcagcttg gctccttcgc    6060

tttctgggct cagaggctgg gaaggggtgg gtccgggggc gggctcaggg gcgggctcag    6120

gggcggggcg ggcgcgaagg tcctccggac ccggcattct gcacgcttca aaagcgcacg    6180

tctgccgcgc tgttctcctc ttcctcatct ccgggccttc gacctgcatg aaaaagcctg    6240

aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc gtctccgacc    6300

tgatgcagct ctcggagggc gaagaatctc gtgctttcag cttcgatgta ggagggcgtg    6360

gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt tatgtttatc    6420

ggcactttgc atcggccgcg ctcccgattc cggaagtgct tgacattggg gaattcagcg    6480

agagcctgac ctattgcatc tcccgccgtg cacagggtgt cacgttgcaa gacctgcctg    6540

aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg atcgctgcgg    6600

ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc ggtcaataca    6660

ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac tggcaaactg    6720

tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg atgctttggg    6780

ccgaggactg ccccg                                                     6795
```

<210> 40
<211> 5821
<212> DNA
<213> Artificial

<220>
<223>   recombinant DNA vector

<400> 40
```
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac      60

tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta     120

aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt     180

taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata     240

gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc     300

agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac     360

cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag     420

tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac     480
```

```
gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc    540

agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg    600

gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc    660

atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct    720

gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc    780

tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc    840

atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc    900

agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc    960

gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca   1020

cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt   1080

tattgtctca tgagcggata catatttgaa tgtatttaga aaataaaca aatagggggtt   1140

ccgcgcacat ttccccgaaa agtgccacct gacgtcgacg gatcgggaga tctcccgatc   1200

ccctatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc cagtatctgc   1260

tccctgcttg tgtgttggag gtcgctgagt agtgcgcgag caaaatttaa gctacaacaa   1320

ggcaaggctt gaccgacaat tgcatgaaga atctgcttag ggttaggcgt tttgcgctgc   1380

ttcgcgatgt acgggccaga tatacgcgtt gacattgatt attgactagt tattaatagt   1440

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta   1500

cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga   1560

cgtatgttcc catagtaacg ccaatgggga ctttccattg acgtcaatgg gtggagtatt   1620

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta   1680

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg   1740

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt   1800

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc   1860

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat   1920

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct   1980

atataagcag agctctctgg ctaactagag aacccactgc ttactggctt atcgaaatta   2040

atacgactca ctatagggag acccaagctg gctagcgttt aaacttaagc ttgccaccat   2100

gaagctactg tcttctatcg aacaagcatg cgatatttgc cgacttaaaa agctcaagtg   2160

ctccaaagaa aaaccgaagt gcgccaagtg tctgaagaac aactgggagt gtcgctactc   2220

tcccaaaacc aaaaggtctc cgctgactag ggcacatctg acagaagtgg aatcaaggct   2280

agaaagactg gaacagctat ttctactgat ttttcctcga agagccttg acatgatttt   2340

gaaaatggat tctttacagg atataaaagc attgttaaca ggattatttg tacaagataa   2400
```

```
tgtgaataaa gatgccgtca cagatagatt ggcttcagtg gagactgata tgcctctaac    2460

attgagacag ttgactgtat cgccgggtac cgagctcgga tccactagtc cagtgtggtg    2520

gaattctgca gatatccagc acagtggcgg ccgctcgagt ctagagggcc cgtttaaacc    2580

cgctgatcag cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctccccc    2640

gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa    2700

attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac    2760

agcaaggggg aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg    2820

gcttctgagg cggaaagaac cagctggggc tctagggggt atccccacgc gccctgtagc    2880

ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc    2940

gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt    3000

ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc tttacggcac    3060

ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc gccctgatag    3120

acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa    3180

actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg gattttgccg    3240

atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc gaattaattc    3300

tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg ctccccagca ggcagaagta    3360

tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg aaagtcccca ggctccccag    3420

caggcagaag tatgcaaagc atgcatctca attagtcagc aaccatagtc ccgcccctaa    3480

ctccgcccat cccgcccta actccgccca gttccgccca ttctccgccc catggctgac    3540

taattttttt tatttatgca gaggccgagg ccgcctctgc ctctgagcta ttccagaagt    3600

agtgaggagg cttttttgga ggcctaggct tttgcaaaaa gctcccggga gcttgtatat    3660

ccattttcgg atctgatcaa gagacaggat gaggatcgtt tcgcatgatt gaacaagatg    3720

gattgcacgc aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac    3780

aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag gggcgcccgg    3840

ttcttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac gaggcagcgc    3900

ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg    3960

aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc ctgtcatctc    4020

accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc    4080

ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta    4140

ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat caggggctcg    4200

cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag gatctcgtcg    4260

tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc ttttctggat    4320
```

```
tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc    4380

gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta    4440

tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag ttcttctgag    4500

cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat cacgagattt    4560

cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc gggacgccgg    4620

ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc ccaacttgtt    4680

tattgcagct tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc    4740

atttttttca ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt    4800

ctgtataccg tcgacctcta gctagagctt ggcgtaatca tggtcatagc tgtttcctgt    4860

gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa    4920

agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc    4980

tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag    5040

aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt    5100

cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga    5160

atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg    5220

taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa    5280

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt    5340

tcccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct    5400

gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct    5460

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc    5520

cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt    5580

atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc    5640

tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat    5700

ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa    5760

acaaaccacc gctggtagcg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa    5820

a                                                                    5821
```

<210> 41
<211> 10075
<212> DNA
<213> Artificial

<220>
<223> recombinant DNA molecule

<220>
<221> misc_feature

<222> (9040)..(9040)
<223> unsequenced nucleotide

<400> 41

```
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc      60

aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt     120

aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa     180

aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat     240

gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct     300

gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg     360

caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag     420

ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta     480

attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg     540

ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg     600

gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct     660

ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta     720

tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg     780

gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc     840

cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg     900

gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga     960

tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg    1020

ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat    1080

gttgaatact catactcttc cttttcaat attattgaag catttatcag ggttattgtc    1140

tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca    1200

catttccccg aaaagtgcca cctaaattgt aagcgttaat attttgttaa aattcgcgtt    1260

aaatttttgt taaatcagct cattttttaa ccaataggcc gaaatcggca aaatccctta    1320

taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga caagagtcc    1380

actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg    1440

cccactacgt gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact    1500

aaatcggaac cctaaaggga gccccgatt tagagcttga cggggaaagc cggcgaacgt    1560

ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc    1620

ggtcacgctg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc    1680

ccattcgcca ttcaggctgc gcaactgttg gaagggcga tcggtgcggg cctcttcgct    1740

attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg    1800
```

```
gttttcccag tcacgacgtt gtaaaacgac ggccagtgag cgcgcgtaat acgactcact    1860

ataggggcgaa ttggagctcc accgcggtgc ggccgggcca tgcaggccac gacatgataa    1920

gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttatt    1980

gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagttc    2040

ttgtacagct cgtccatgcc gagagtgatc ccggcggcgg tcacgaactc cagcaggacc    2100

atgtgatcgc gcttctcgtt ggggtctttg ctcagggcgg actggtagct caggtagtgg    2160

ttgtcgggca gcagcacggg gccgtcgccg atggggggtgt tctgctggta gtggtcggcg    2220

agctgcacgc tgccgtcctc gatgttgtgg cggatcttga agttcacctt gatgccgttc    2280

ttctgcttgt cggccatgat atagacgttg tggctgttgt agttgtactc cagcttgtgc    2340

cccaggatgt tgccgtcctc cttgaagtcg atgcccttca gctcgatgcg gttcaccagg    2400

gtgtcgccct cgaacttcac ctcggcgcgg gtcttgtagt tgccgtcgtc cttgaagaag    2460

atggtgcgct cctggacgta gccttcgggc atggcggact tgaagaagtc gtgctgcttc    2520

atgtggtcgg ggtagcgggc gaagcactgc aggccgtagc cgaaggtggt cacgagggtg    2580

ggccagggca cgggcagctt gccggtggtg cagatgaact tcagggtcag cttgccgtag    2640

gtggcatcgc cctcgccctc gccggacacg ctgaacttgt ggccgtttac gtcgccgtcc    2700

agctcgacca ggatgggcac cacccccggtg aacagctcct cgcccttgct caccatggtg    2760

gcgttagctt gatttgacag tggctggggg ttgcgccgcc gggttttata ggaagccaca    2820

gcggccactc gagccataaa aggcaacttt aggaacggcg gggggtgatt ggattcgagt    2880

cgtttattca ccggccttgc cgcacagtgc agcattttt tacccccctct cccctccttt    2940

tgcggggggaa aaaaaaaaa aaaaaaaaaa aaggagaaga gaaaaaaagc gagcgagaga    3000

gaaagcgaga ttgaggaaga ggatgaagag tttggcgatg ggtgctggtt ccgtaggccc    3060

agatggacaa gaatagcccc cgcccttgcg gacagtatcc cattcagtga ctcagatcag    3120

atcaagcggc cgccagtgtg atggatatct gcagaattcc agcacactgg cggccgttac    3180

tagtggatcc ctctggagct gcctagtgtg ggtgtctgct ggaaacttct aaaatgtctc    3240

ggcctatta attggagaat tcactgaaga attaaaggac tcagtaaaga agaggtttgg    3300

tttgttcatg aaagtaagta gcttgtgaaa tgctttattt tcctgtgact aaaattccac    3360

cttgtaattc ccaataagga ttgcttcaag ctatttaata caaattaaac ctgtttcacc    3420

atgctcatag gaaaatgtca aagtatcttc attcaccaac tgagggctgg ctcttctttg    3480

aaggtagctc tgagatttgt tttgtaatgg agacaaagac ttatttaact gttggggagg    3540

gggaaggagg acttaggcca ggcaggcggg gatgtgggggg acgaaccatg aggggcaatt    3600

tggctgcctc tgaaaaggtc cagagctacc ttaaaagaca ggctttctat aacaagtggt    3660

cctttctttt cctttctttt ttttttttgag acagagtctc actctgttgc ccaagctgga    3720
```

```
gtgcagtggt gcgatctcag ctcactgcaa cctccagctc ccaggttcaa gcaattctcc    3780

tgcctcagcc tcccaagtag ctgggattac aggcaaccgc taccatgccc agctaatttt    3840

ttgtattttt aatagagatg gggtttcacc atgttggcca gggtggtctc aaactcctga    3900

cctcaagtga tccacccacc tcagcctccc aaagtgctgg gattacagac atgagccacc    3960

gcgcccagac acagctggct ttgtcgccca ggctggaagt gcagtggcac gatctcggct    4020

cactgcaagc tctgcctcct gggttcacgc cattctcctg cctcagcttc ccaagtagct    4080

gagactacag gcgcccgcca ccacgcccgg ctaatttttt gtatttttag tggagatggt    4140

gtttcaccat gttagccagg atggtcttga tctcctgacc tcgtgatccg cctgtcttgg    4200

cctcccaaag tgctgggatg acaggcgtga gccaccacgc ctggcttaca agtggccctt    4260

tctaagcccc aagatgtcac tgtctttca ctgaaggcgg aaaaagaatg taggcatgtt     4320

tgtctgagag ggagggagta gtttaggag aaaggaccaa gagcacattt tgagtgggaa      4380

gtaagtggtt cctttcta gccttatgat atggtttggc tctgtgtccc cacccaaatc       4440

tcatggtgaa ctgtaatccc cagtggtgga ggtgaggcct ggcgggaggg actggatcat    4500

gggggtgatc cttcatgaat ggtttagcac catcctcttg gtgctactct agtgatagcg    4560

ttctcacaag atctgatggg tgaaaagtgt gtgtctcttg tttttgctcc tgccatgcaa    4620

gacgcctact cccacttcac cttctgctat gagtcaaatt ccctgaggcc tccccagaag    4680

cagaagcttc ctgtacagcc ttgtgggcca cgagccaatt aaacctcttt tctttataaa    4740

ttaccaggtc tcagggattt ctttatagca gactaataca cctcaaagta gggatttatt    4800

ttattattct attttttttt tgagataggg tctcgatctg tcacccaggc tggagtgcag    4860

tggcatgatc tcagcttgct gtaacctctg cctcccgggt tgaagcaatt ctcatgccta    4920

agccacccaa gtagctggaa ttacaggtgt gtgccaccct gctcagctaa tttttgtatc    4980

tttagtagag ttggggtttc tccatgttgg ccaggctggt atcaaactcc tggactcaag    5040

caatctgccc accttggcct ccccagagtg ccaggattac aggcatgacc catcttgcct    5100

ggtcagtttt ttttttttta aatccagtta atgctgggtg ccatgactca cacctgtaat    5160

gccagtgctt tgggaggccg aggtgggaag atcacttgag gccaggagtt tgagaccagc    5220

ctgggaaaca tctcatctct accaaaaaac aaaaacaaaa caaacaaca acaacatcaa      5280

acatagccag gtgtggtggt gcacacctgt agtcccagct ccttgggagg ctgaggcagg    5340

aggattgctt gagcccagga gttcaaggct gcagtgaact atgattgcac cactgcattc    5400

cagcctgggc aacagagcaa gaacctatct ctatttaaga aaaatgctc ctaagaaatc      5460

atcttagtac ttggtaagct attaaggttt tttttaagtg tattaattac atagccagta    5520

taataactta gttgctaaga gctttgaggt ggctccaagc aagaagaaag gagttgttgt     5580

taagctacag ctcagagtct cagaatgaac tgtttccatc ttcgctgtcc tttgcttaac    5640
```

```
tgtgggataa atataatttt ctcaacattc aacacatgtg tccttaatcg tcctctacga    5700

cactggttcc atgttgccaa gagaacccag aacagccgca aacacctact ggattgttac    5760

tggcctttcc tgagaggatg acccgggctt tgaccttgtt catgttgaag ttcttcaagt    5820

aggcctcgta gattctcttg gccagagatt tgagatctgc agtttcagaa tcttctatgt    5880

catgttcaca ggtaagaatt tctgctttca gttttgcttt ctcagatctt ggcattcgtc    5940

cggtacccgg cgatacagtc aactgtcttt gacctttgtt actactctct tccgatgatg    6000

atgtcgcact tattctatgc tgtctcaatg ttagaggcat atcagtctcc actgaagcca    6060

atctatctgt gacggcatct ttattcacat tatcttgtac aaataatcct gttaacaatg    6120

cttttatatc ctgtaaagaa tccattttca aaatcatgtc aaggtcttct cgaggaaaaa    6180

tcagtagaaa tagctgttcc agtctttcta gccttgattc cacttctgtc agatgtgccc    6240

tagtcagcgg agaccttttg gttttgggag agtagcgaca ctcccagttg ttcttcagac    6300

acttggcgca cttcggtttt tctttggagc acttgagctt tttaagtcgg caaatatcgc    6360

atgcttgttc gatagaagac agtagcttca tggtggcaag cttaagttta aacgctagaa    6420

ctagtggatc cggataagcc agtaagcagt gggttctcta gttagccaga gagctctgct    6480

tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg cggagttgtt    6540

acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat tgacgtcaat    6600

ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat tgatgtactg    6660

ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact gccaagtagg    6720

aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac cgtcattgac    6780

gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt gggcagttta    6840

ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta ctatgggaac    6900

atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag gcgggccatt    6960

taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg accccgtaat    7020

tgattactat taataactag tcaataatca atgtcaaccc cgggctgcag gaattctacc    7080

gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg    7140

cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc    7200

caaccggctc ccttctttgg tggccccttc gcgccacctt ctactcctcc cctagtcagg    7260

aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt    7320

ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg    7380

gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg    7440

tgggtccggg ggcgggctca ggggcgggct caggggcggg gcggcgcga aggtcctccg    7500

gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca    7560
```

```
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt    7620

ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat    7680

ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg    7740

ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga    7800

ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc    7860

gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc    7920

cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg    7980

gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga    8040

ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg    8100

tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc    8160

tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg    8220

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct    8280

tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc    8340

cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac    8400

tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg    8460

acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg    8520

cggccgtctg accgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca    8580

gcactcgtcc gagggcaaag gaatgcctga aaaggaagt gagctgtaaa ggctgagctc    8640

tctctctgac gtatgtagcc tctggttagc ttcgtcactc actgttcttg actcagcatg    8700

gcaatctgat gaaatcccag ctgtaagtct gcataaattg atgatctatt aaacaataaa    8760

gatgtccact aaaatggaag ttttttactg tcatactttg ttaagaaggg tgagaacaga    8820

gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg gtgggattag    8880

ataaatgcct gctctttact gaaggctctt tactattgct ttatgataat gtttcatagt    8940

tggatatcat aatttaaaca agcaaaacca aattaagggc cagctcattc ctccactcat    9000

gatctataga tctatagatc tctcgtggga tcattgtttn tctgatccac tggaagctta    9060

tcgataccgt cgacctcgag gggggcccg gtaccagct tttgttccct ttagtgaggg    9120

ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg    9180

ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa    9240

tgagtgagct aactcacatt aattgcgttg cgctcactgc cgctttcca gtcgggaaac    9300

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt    9360

gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg ctgcggcga    9420

gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca    9480
```

ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg   9540

ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt   9600

cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc   9660

ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct   9720

tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc   9780

gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta   9840

tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca   9900

gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag   9960

tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag   10020

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccg   10075

<210>   42
<211>   8971
<212>   DNA
<213>   Artificial

<220>
<223>   recombinant DNA vector

<400>   42
gtaggtgtca ttctattctg ggggtgggg tggggcagga cagcaagggg gaggattggg   60

aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaaagaa   120

ccagctgggg ctctaggggg tatccccacg cgccctgtag cggcgcatta agcgcggcgg   180

gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt   240

tcgctttctt cccttccttt ctcgccacgt tcgccggctt ccccgtcaa gctctaaatc   300

ggggctccc ttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg   360

attagggtga tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga   420

cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc   480

ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa   540

aaaatgagct gatttaacaa aaatttaacg cgaattaatt ctgtggaatg tgtgtcagtt   600

agggtgtgga agtccccag gctccccagc aggcagaagt atgcaaagca tgcatctcaa   660

ttagtcagca accaggtgtg gaaagtcccc aggctcccca gcaggcagaa gtatgcaaag   720

catgcatctc aattagtcag caaccatagt cccgccccta actccgccca tcccgcccct   780

aactccgccc agttccgccc attctccgcc ccatggctga ctaattttt ttatttatgc   840

agaggccgag gccgcctctg cctctgagct attccagaag tagtgaggag gcttttttgg   900

aggcctaggc ttttgcaaaa agctcccggg agcttgtata tccattttcg gatctgatca   960

gcacgtgatg aaaaagcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa   1020

```
gttcgacagc gtctccgacc tgatgcagct ctcggagggc gaagaatctc gtgctttcag      1080

cttcgatgta ggagggcgtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta      1140

caaagatcgt tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgct      1200

tgacattggg gaattcagcg agagcctgac ctattgcatc tcccgccgtg cacagggtgt      1260

cacgttgcaa gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc      1320

catggatgcg atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc      1380

gcaaggaatc ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca      1440

tgtgtatcac tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct      1500

cgatgagctg atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga      1560

tttcggctcc aacaatgtcc tgacggacaa tggccgcata acagcggtca ttgactggag      1620

cgaggcgatg ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg      1680

gttggcttgt atggagcagc agacgcgcta cttcgagcgg aggcatccgg agcttgcagg      1740

atcgccgcgg ctccgggcgt atatgctccg cattggtctt gaccaactct atcagagctt      1800

ggttgacggc aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg      1860

atccggagcc gggactgtcg ggcgtacaca aatcgcccgc agaagcgcgg ccgtctggac      1920

cgatggctgt gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcgtccgag      1980

ggcaaaggaa tagcacgtgc tacgagattt cgattccacc gccgccttct atgaaaggtt      2040

gggcttcgga atcgttttcc gggacgccgg ctggatgatc ctccagcgcg gggatctcat      2100

gctggagttc ttcgcccacc ccaacttgtt tattgcagct tataatggtt acaaataaag      2160

caatagcatc acaaatttca caaataaagc attttttca ctgcattcta gttgtggttt       2220

gtccaaactc atcaatgtat cttatcatgt ctgtataccg tcgacctcta gctagagctt      2280

ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca      2340

caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact      2400

cacattaatt gcgttgcgct cactgcccgc tttccagtcg gaaacctgt cgtgccagct       2460

gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc      2520

ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca      2580

ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg      2640

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttttcca     2700

taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa       2760

cccgacagga ctataaagat accaggcgtt ccccctgga agctccctcg tgcgctctcc       2820

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc      2880

gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct      2940
```

```
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg    3000

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag    3060

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta    3120

cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg    3180

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gttttttttgt    3240

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc    3300

tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt    3360

atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta    3420

aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat    3480

ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac    3540

tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg    3600

ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag    3660

tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt    3720

aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt    3780

gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt    3840

tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt    3900

cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct    3960

tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt    4020

ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac    4080

cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa    4140

actctcaagg atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa    4200

ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca    4260

aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct    4320

ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga    4380

atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc    4440

tgacgtcgac ggatcgggag atctcccgat cccctatggt gcactctcag tacaatctgc    4500

tctgatgccg catagttaag ccagtatctg ctccctgctt gtgtgttgga ggtcgctgag    4560

tagtgcgcga gcaaaattta agctacaaca aggcaaggct tgaccgacaa ttgcatgaag    4620

aatctgctta gggttaggcg ttttgcgctg cttcgcgatg tacgggccag atatacgcgt    4680

tgacattgat tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc    4740

ccatatatgg agttccgcgt tacataactt acggtaaatg cccgcctgg ctgaccgccc    4800

aacgaccccc gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg    4860
```

```
actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat    4920

caagtgtatc atatgccaag tacgcccct attgacgtca atgacggtaa atggcccgcc     4980

tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta    5040

ttagtcatcg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag    5100

cggtttgact cacggggatt tccaagtctc caccccattg acgtcaatgg gagtttgttt    5160

tggcaccaaa atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa    5220

atgggcggta ggcgtgtacg gtgggaggtc tatataagca gagctctctg gctaactaga    5280

gaacccactg cttactggct tatcgaaatt aatacgactc actataggga gacccaagct    5340

ggctagcgtt taaacttaag cttgccacca tgaagctact gtcttctatc gaacaagcat    5400

gcgatatttg ccgacttaaa aagctcaagt gctccaaaga aaaaccgaag tgcgccaagt    5460

gtctgaagaa caactgggag tgtcgctact ctcccaaaac caaaaggtct ccgctgacta    5520

gggcacatct gacagaagtg gaatcaaggc tagaaagact ggaacagcta tttctactga    5580

ttttcctcg agaagacctt gacatgattt tgaaaatgga ttctttacag gatataaaag    5640

cattgttaac aggattattt gtacaagata atgtgaataa agatgccgtc acagatagat    5700

tggcttcagt ggagactgat atgcctctaa cattgagaca gcatagaata agtgcgacat    5760

catcatcgga agagagtagt aacaaaggtc aaagacagtt gactgtatcg ccgggtaccg    5820

cagtggggat gtctcataat ggtaagtaaa cagtcatcac catatacttt attattctca    5880

ttatagctgc cagaccagtg gacactaaag ccattgccaa aaatgtgtac agttttcca     5940

ccaaatgcca gaatttagaa tattgcatgg cgataaaaca tttctctttt aggtcagtgt.    6000

ttttaaagtt ttattataga acctttctct ctgtggttgg gcatctgcca tgaggagaaa    6060

agagacttga aaaatctggg ggattatggg aaaaacttta gtgttagaac aaatataatg    6120

atcatcatga ttaattttta tattgcctca tgcattaaaa acagaattaa aatttacttt    6180

tgtgtatcag tagtgtcatt ctgctttaaa tgtaagaaaa tcctgaagat gtctgggtct    6240

gaaactttaa atggcatatg agaatttaga ttcatttaat ttttttttt ttttttttt     6300

tgctgtttat tatgcaaaga cgagtattcg tttccttcac ttcaaggaac gtccatgagt    6360

tgtggaaagg gtaaaacaaa ttatttttat agttctgcaa aacttcaaag ccatagaaac    6420

ataagagaat aaaactttct acttttgaaa cctgtgtcta aagtcaagct tatgactttg    6480

ggtgtcaact taagggaaat gatgtagaag acctgagaaa agacccatgt tttcatgatc    6540

tgtggaacta actcagtaag gtcctggaag acaaacgtgt ttgaaaaaat aattcttggt    6600

cactatttgg ttctgattgg ttgcctggct gacaaaataa attttaggtg taacttgata    6660

gtattatatg tgtctagtat taatcagtta ttttcaatgc actcaaggaa taaaaaatgc    6720

caacatgatc aaattaaatt tgaagtttca tttcatcagc tatcctattt tttctttctt    6780
```

```
ttgttttcat ctctctcaaa actgatctag taagaaggca gtcttcatct atatgtactc   6840

taaagcatct gagaaaatta aatttttgct ctatcatgca gatacttggt tttcacagtg   6900

gaaaccatca agatatttga gagtcactca ttgattaata acctaccctt taaaaatagt   6960

tcacaaaatt ctatctgtat gattactaat gggtcagttc ctctaaaggc aagaatgtat   7020

tttgaataga gctatcagca gtactgaaca ggatggaaat ttggtccaac tttcctcaac   7080

aattgactac atctgctgta tacaactgta gtcatggatt gcaaataat ttgctggttc   7140

agttccatca gtttcctttt tatttcagac ctcataaaac ataagggcaa gcttccaagt   7200

caagggttgt attttaggtg atgcaatcaa gttgcaataa acttatttca tcttcatgaa   7260

gaaatcattg actttgaggt cattatcact gagacagcct aggaaagcat gccagttccg   7320

cttttaacac taattttgc ctattaaaaa attatcactt cacacattca tgtaaaatat   7380

aattcctgtt acttacaggc agtgcataca ttgatgtaaa atatatttcc tgttgcttac   7440

aggcattgca tacgtttact taaaatcaat ctgcaattgc ttattgttta taggcaaata   7500

ttatttaaac cccaacctat gtagtatgga tggtataatt ttctcctccc ataaatgtca   7560

aagaaagcta tcataaactt gataggtcat attttgactt catatgacaa tttattttat   7620

ttctagagag ggtcaatctt gatagggagt gtttgttcag ttctagcaaa agcaagaaaa   7680

tatatgaatt aagtgtcatg acatcagaga cttaggcagt tcaacactta aaaaattgca   7740

aacataacta agaaaacctc agaggttttt atgggagttt gtttgctgca cagaaaataa   7800

acatcagact aacaacaagg ttttgattag ttttgcctgt tagctgtatg aagaaaacat   7860

ttctaaatac aaagttttct aggaagttga gtcactattt aaaatccatt gctttgacca   7920

atgtgaagcc aacagggttt tcattcatat agaatgcctt tttatcaaaa gaggctctga   7980

gtgttatacc aaaatcagca cccagctgtt ttcagtatgc gcaagaaaac atgaacgtcc   8040

cccaaagcaa atgtactcca ttccgtgtta accttatttc tgctccctgg tctctgatta   8100

gtctggttga tgaatgtagc ttcccctttg gccatgcact cacagctttg aaggaggctt   8160

ctgaaggaac tggtagggaa ccctaggggc tcttgctaat gttgttggat agtcactgaa   8220

atcagtgatc accatgtcct ttcactttga aaattctatg attctgtata atcaaacaac   8280

ctaactcact ggatttttca cttttcaacc agctattccg tgtcttacag tgctctggtc   8340

ttcacagata ggtaggcacc tggcatgtca gttgcctgaa tttgaaagtt ttcacctgta   8400

tgttttggta cgataaaaat aaaaatgtaa tttatatatc tgaatcaggt ctgtatgtta   8460

tgatcaattg ctcagcaatt tcgggcagtt ggtttgatgg ttatgtagta atgtagcctg   8520

agagcagaaa tacagagcct ctgggctaga gaaagtataa atggcatcct aggctatgta   8580

gggttacagc tcttcagaag gaactttcat tttcattgtg acacatcgtc tacatgttgt   8640

agaagaacat agtttcagaa ttcttccagt tagaaacata gtttctcaaa tattcacttt   8700
```

cagcattggg tagaaaaagt acccaggtga tggactgatt gaggatccac tagtaacggc   8760

cgccagtgtg ctggaattct gcagatatcc atcacactgg cggccgctcg agtctagagg   8820

gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt   8880

ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta   8940

ataaaatgag gaaattgcat cgcattgtct g   8971

<210> 43
<211> 29
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 43
cgggtccccg gcgatacagt caactgtct   29

<210> 44
<211> 30
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 44
taaagcttgc caccatgaag ctactgtctt   30

<210> 45
<211> 10514
<212> DNA
<213> Artificial

<220>
<223> recombinant DNA vector

<400> 45
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aattttttgtt aaatcagctc   60

attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga   120

gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc   180

caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc   240

ctaatcaagt ttttggggt cgaggtgccg taaagcacta atcggaacc ctaaagggag   300

cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa   360

agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac   420

cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg   480

caactgttgg gaaggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg   540

gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg   600

```
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca        660

ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca        720

aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc        780

tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt        840

ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct        900

tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca        960

gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc       1020

cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg       1080

ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc       1140

cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga       1200

acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct       1260

ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt       1320

agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg       1380

gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct       1440

cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga       1500

tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat       1560

ttgacagtgg ctggggggttg cgccgccggg ttttatagga agccacagcg gccactcgag       1620

ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg       1680

gccttgccgc acagtgcagc attttttttac cccctctccc ctccttttgc gggggaaaaa       1740

aaaaaaaaaa aaaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg       1800

aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa       1860

tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg       1920

tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat       1980

ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc       2040

agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga       2100

ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac       2160

aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc       2220

aggagaattg cttgaactcg ggaagcagag gttgcagtga ccaagatca tgccattgca       2280

ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaacaa aaacaaaaaa       2340

acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt       2400

tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca       2460

caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg       2520
```

```
tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt    2580

agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt    2640

gcaaagtttc aatttttaat ctattgattt gaaagaatta ttatacatta ttttaaatat    2700

cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat    2760

taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt    2820

agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag    2880

agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa    2940

tctgtgggca catatttaaa ctctcaataa atattaactt tgattatta aaaaatacct     3000

ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa    3060

atgagaaaat tttcttctgc actgaattct tcttttaact aaattttaag aaaaaaacca    3120

atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag    3180

gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt    3240

accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt    3300

ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa    3360

attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc    3420

agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga    3480

atccataata aaatgtagaa gagtctgttg atctggggtg agttcagttt tctcctacac    3540

tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca    3600

ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg    3660

aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc    3720

acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat    3780

ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt    3840

cactgaaagt tcaagcttta attataacta atgaaaggtt ttcatatagt atgactacta    3900

tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt    3960

attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggattttat    4020

aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctacttttgc    4080

actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cctttaaaga    4140

aaaaagttt gtcaacccct aaaatagaga gtcaacaatc tggacaattc agaccccaaa     4200

catctctttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg    4260

ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatataaa    4320

atacttttc ttaaataaat acaaaaatac gtatttgtaa aactaaatgt ggataaagtt     4380

ttctttgtca ataatgtaat aaatattgag aagggtgtgg tttggacaat catagcttac    4440
```

```
tggttttttca gctaatggca aaataaaatt ccatcaaatg tcctgtgcat tatcacaatg    4500

tgtgtgtcta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt    4560

tcttgagaaa tttgtaaccc cagtgggtag gactttctc tattcatcca catacgtgta     4620

ttcaccctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct    4680

tcccagtttg agactgtctt ctcaaaaga gggaaaaaat ggtcatcaat actcagttgg     4740

aaaaaatcca ttttatacac caaatcattt ctgctggcaa acactaggta gtaggactgg    4800

aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa    4860

tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca    4920

gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact    4980

cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt    5040

cgaggtcact tgtcgcaagt cacgaccttc actgtcttct ctagacggcg atacagtcaa    5100

ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg    5160

tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt    5220

attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatcct gtaaagaatc    5280

cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag    5340

tctttctagc cttgattcca cttctgtcag atgtgccta gtcagcggag accttttggt      5400

tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgcact tcggttttc     5460

tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag    5520

tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga    5580

gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg    5640

cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat    5700

tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat    5760

tgatgtactg ccaaaacgc atcaccatgg taatagcgat gactaatacg tagatgtact     5820

gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac    5880

cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt    5940

gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta    6000

ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag    6060

gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg    6120

accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg    6180

cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc    6240

gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg    6300

cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc    6360
```

```
caaccggctc ccttctttgg tggcccccttc gcgccacctt ctactcctcc cctagtcagg      6420

aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt      6480

ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg      6540

gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg      6600

tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg      6660

gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca      6720

tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt      6780

ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat      6840

ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg      6900

ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga      6960

ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc      7020

gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc      7080

cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg      7140

gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga      7200

ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg      7260

tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc      7320

tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg      7380

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct      7440

tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc      7500

cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac      7560

tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg      7620

acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg      7680

cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca      7740

gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa      7800

agatgtccac taaaatggaa gtttttcctg tcatactttg ttaagaaggg tgagaacaga      7860

gtacctacat tttgaatgga aggattggag ctacgggggt ggggtgggg tgggattaga      7920

taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt      7980

ggatatcata atttaaacaa gcaaaccaa attaagggcc agctcattcc tcccactcat      8040

gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt      8100

ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc      8160

agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca      8220

gatttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc      8280
```

```
gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt      8340

aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct      8400

cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg      8460

agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct      8520

gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg      8580

gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc      8640

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg     ·8700

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct      8760

ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca      8820

gaggtggcga aacccgacag gactataaag ataccaggcg tttcccctg gaagctccct       8880

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc      8940

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt      9000

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc      9060

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc      9120

cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg      9180

gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc      9240

agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag      9300

cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga       9360

tcctttgatc ttttctacgg ggtctgacgc ·tcagtggaac gaaaactcac gttaagggat      9420

tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag      9480

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat      9540

cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc      9600

cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat      9660

accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag      9720

ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg      9780

·ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc      9840

tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca      9900

acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg      9960

tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc     10020

actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta     10080

ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc     10140

aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg     10200
```

```
ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc    10260

cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc    10320

aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat    10380

actcatactc ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag    10440

cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc    10500

ccgaaaagtg ccac                                                       10514
```

<210> 46
<211> 9626
<212> DNA
<213> Artificial

<220>
<223> recombinant DNA vector

<400> 46
```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg     60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420

attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt    480

atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540

atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660

actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720

aaaatcaacg ggactttcca aaatgtcgta caactccgc cccattgacg caaatgggcg    780

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca    840

ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc    900

gtttaaactt aagcttggta ccgagctcgg atccactagt ccagtgtggt ggaattctgc    960

agatatccag cacagtggct tgcggccgcg aaggaagata gccaggctt actgaggctg    1020

gtggagggag ccactgctgg gctcaccatg accgccgga tgtggggggc ccacgtcttc    1080

tgcgtgttga gcccgttacc gaccgtgagt agccagctga gacccctctgg cctctgccac    1140

tgggctccag atttgggagc cacagggcca actggagcct ggcggagacc ctactgccct    1200

tctcctgctc tagagtccct ccctacgagc aggtggcagg ctgtggcttg gagagcccct    1260
```

```
gacactggggg ctccctccct catctccctc tgcccagcaa tctgtagtga ctctgggtag    1320

tgctattaac atgtccttcc aggggctgag ctgagctcac cttcttgttg cttcccccag    1380

ggcctggctg tgcctctggg ggccacagag actctgtctg ccacaagaca gtgcctcaga    1440

attggtgagg caaagtcagg ctcccctctt agtccctgca gtttgctctt tctctcagct    1500

gtgcacctgt gatgcttcca cacgctcctg tggtaagctt tctggaccct tccttagccc    1560

tatgactccc tgagccactc tccagtttgg tggtttcctt tgaagagggt gtggccaggg    1620

aaggaatgga acctaaggag ttggtggaga ccagctgctt ctgctccagc acttagggga    1680

tccatcatct tctcacgatg ctgagggctg gagaggcttt tggcaaccaa attgtactcc    1740

atgggctaga gcgctggagc acggggcatc aggagagagg gcccttcctg aatgggctgc    1800

agtttcccag taactttcca ggaggttact gcaagtgaca gcagggacca ggctgggagt    1860

cacaaggttc tgggatggtg agggcctgaa gcattcttcc tccgctgcct tattggaggg    1920

tgctgcccct cccgagcacc ttgctgcacc tgctcctgag tgcacaggct ttgagcttga    1980

atcctggacg catgaaacta gccgtgtgaa ctgcaagttc ctcagccttc ctgagtcttg    2040

cttttatctc ttttgttaaa agagaaaata tccgctggcc cgggtcactg tgcattgagt    2100

ggaatctgat ttgtcaaggg ctgggcatag gggtgcaaca caggaaactt cgactgcccc    2160

cctgtctgcc atctgtgtgc actctacagc aaaagcttta tttttagcag gggtacccaa    2220

gatctggtgg tggccattca gcagatggtg ctagtaccaa agctagaact caatcactgt    2280

ttgccagtgt gggcatcatc ccctgcagat accaggaaga gcagcaggca tagatgcaag    2340

gcctgggggt ccagggcgat ggagcctccc ctcaggggtg ggtggggtcc acgatgtgtg    2400

tgggggggtg actgggagcc ccaggatgtg tttcgccagt cctaattcag caggagcatt    2460

tggataaatc ttagaacgat ctatccttcc actcccacca cttccctccc tagaatctct    2520

ccttacatct ggccctatgc ccacacatcc tgagtcatta gtggggcttc tctgaatttc    2580

cctccttctc ttagaggtgt gggcactgtc tgttcatcct tttggcatga ggctgctctc    2640

ctggcatcct ttgccctcag ttgcttcccc aaattggccc tgaaaagttt cccaggcagg    2700

gttggggctt gccagactgt gcagtaggca gaaggatagg cagagaccac agctcttcca    2760

acagcccagg gtcaatatca ggctgatggt gtgagaaagt cccatacagg cccatctgcc    2820

ctcctgctct tccttccccg ctttgtgcgc tcagtctgtt gggagatgta ggggagacgc    2880

aacggggctc tgccctgaga tggacctgag gttactgagt cctcgggagg aagccaagag    2940

ctgagtccgg ggcacccacc aaaataagac ctggagcctg ggtagtgggt gacatctgca    3000

ccctgggggcc ttgcagggcc aggcaggaga tggcagagct gcgtccaagc catggaccca    3060

tggggtcagc agggcgggac aggagtgact tctcctgcca gggtagcact ggtggggggtg    3120

accctcctta ctgggctggc tctccagatt ttctcctatc ccacaaatct gcggcagagc    3180
```

```
tctcgggtct ctccacgtgg tcacccacca tctattgtcc ccactctggt ggtcaaagtc    3240

tctactctct gaccccagcc ctggtcttac atctgctaca ctatttgctg tgcctgcatt    3300

tctcccaaag aacactagct gttaattaaa catttgctga catggatcct gctagacatt    3360

gagggtgtgt gtgtgtgggg tgggggtggg ggccatgggg agagaaacat tagtcctgag    3420

ctgatacttc ccaaatacac cttttcccac ttcccttgtc cttgcacatg ctgtgctttc    3480

tgttccagat gctcttctcc ccactgtgtg tgtctggctc taaaaaatgc tcccttgtcc    3540

ccaggtccag ctcaaatgcc acctccccaa tccctgtgag agggacattc cctcccatgg    3600

tacctgtctt gcttcttttt gttttttttt tgtttttctg agatggagtc tcactttatc    3660

gcccaggctg gagtacagtg gcacaatctc agctcactgt aacctccgcc tcctgggttc    3720

aagtgattct cctgcctcag cctcctgagt agctgggatt acaggagcgt gtcatcacac    3780

ccagctaatt tttgtatttt tgtagagata gtgttttgcc acattggtca ggctggtgtt    3840

taactcctga cgtcaggtga tccgcctacc ttggcctccc aaagtgctgg gattacaggc    3900

atgagccact gcgcccggcc cctgtcctgc ttctgatcac agtttgcgca gggtctatta    3960

ggatccagcc cttctccctt ccccagcccc tcctctcctg tgaaagcctg ggggaacaga    4020

catctacact ggttgacttg gatgtgtgtt ttatacaaaa atgtgtatgc tgtataccca    4080

gttcacagtg cagaaagtat gttttgtaaa cttggcttcc cttgtatgta tctgtttggg    4140

ctctggcctg ctgggcgatc agttttggtt ttccttcggt ttggggcggt gggagggaga    4200

agtgctagaa attcctgctg ttcacttcca gtttcctgga gaacttgtta aaatacaaag    4260

tgctggatcc accctagatc tctacatatt cctgcttcca tcccaatgga tctctagctg    4320

gttcttgagg cacactttca gaagtactgc tagaagattc ccaaaaaaac cccaacttgt    4380

ttccagcttc cagagatcat ttctcatggc cactgtgtag cacaagatct gggaggctcc    4440

taaaactata atctttcaag gttgctcagc catcaaggtt atgatccttt catctacaca    4500

tccatgaagc cacacatcca tccttacaca gctctctcac tcatccacca tgcacccatc    4560

catgcagcca gtcatctgtc tatccatcca tccacctccc agccatata tttgtctgtc    4620

tgtctgtata ctcatccatt ccactagcct tccttccttc ttttctgcct tcctgccttc    4680

ctgccttcct tccctccttc cttttttat gcagccttca cctcattcat gtgctcactc    4740

atccatgtgc tcattcattc atccatcttc tctctcactt actaagctac ccatttgctc    4800

atgtgttcct gcctgtgtgc acccctaaaa gcagagactg agacaaaggt cagaaattta    4860

aaatggtttc cagtggggta aatccaaagt attggcagag ctgcattcct cctggagact    4920

gtagagaata atccatttcc ttgccttttc cttgggttca tgccctttct ccatcttcag    4980

accagcagtg tggcatcttt aagcggccgc tcgagtctag agggcccgtt taaacccgct    5040

gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc tcccccgtgc    5100
```

```
cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg    5160

catcgcattg tctgagtagg tgtcattcta ttctggggg tggggtgggg caggacagca     5220

aggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt     5280

ctgaggcgga aagaaccagc tggggctcta ggggtatcc ccacgcgccc tgtagcggcg     5340

cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc    5400

tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc    5460

gtcaagctct aaatcggggc atcccttag ggttccgatt tagtgcttta cggcacctcg     5520

accccaaaaa acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg    5580

tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg    5640

gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt ttggggattt    5700

cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg    5760

gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc ccaggcaggc agaagtatgc    5820

aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag    5880

gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc    5940

cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa    6000

tttttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc cagaagtagt     6060

gaggaggctt ttttggaggc ctaggctttt gcaaaaagct cccgggagct tgtatatcca    6120

ttttcggatc tgatcagcac gtgatgaaaa agcctgaact caccgcgacg tctgtcgaga    6180

agtttctgat cgaaaagttc gacagcgtct ccgacctgat gcagctctcg gagggcgaag    6240

aatctcgtgc tttcagcttc gatgtaggag ggcgtggata tgtcctgcgg gtaaatagct    6300

gcgccgatgg tttctacaaa gatcgttatg tttatcggca ctttgcatcg gccgcgctcc    6360

cgattccgga agtgcttgac attggggaat tcagcgagag cctgacctat tgcatctccc    6420

gccgtgcaca gggtgtcacg ttgcaagacc tgcctgaaac cgaactgccc gctgttctgc    6480

agccggtcgc ggaggccatg gatgcgatcg ctgcggccga tcttagccag acgagcgggt    6540

tcggcccatt cggaccgcaa ggaatcggtc aatacactac atggcgtgat ttcatatgcg    6600

cgattgctga tccccatgtg tatcactggc aaactgtgat ggacgacacc gtcagtgcgt    6660

ccgtcgcgca ggctctcgat gagctgatgc tttgggccga ggactgcccc gaagtccggc    6720

acctcgtgca cgcggatttc ggctccaaca atgtcctgac ggacaatggc cgcataacag    6780

cggtcattga ctggagcgag gcgatgttcg ggattccca atacgaggtc gccaacatct      6840

tcttctggag gccgtggttg gcttgtatgg agcagcagac gcgctacttc gagcggaggc    6900

atccggagct tgcaggatcg ccgcggctcc gggcgtatat gctccgcatt ggtcttgacc    6960

aactctatca gagcttggtt gacggcaatt tcgatgatgc agcttgggcg cagggtcgat    7020
```

```
gcgacgcaat cgtccgatcc ggagccggga ctgtcgggcg tacacaaatc gcccgcagaa     7080

gcgcggccgt ctggaccgat ggctgtgtag aagtactcgc cgatagtgga aaccgacgcc     7140

ccagcactcg tccgagggca aaggaatagc acgtgctacg agatttcgat tccaccgccg     7200

ccttctatga aaggttgggc ttcggaatcg ttttccggga cgccggctgg atgatcctcc     7260

agcgcgggga tctcatgctg gagttcttcg cccacccaa cttgtttatt gcagcttata     7320

atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt ttttcactgc     7380

attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgt ataccgtcga     7440

cctctagcta gagcttggcg taatcatggt catagctgtt tcctgtgtga aattgttatc     7500

cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct     7560

aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa     7620

acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta     7680

ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc     7740

gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg     7800

caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt     7860

tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa     7920

gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct     7980

ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc     8040

cttcgggaag cgtggcgctt tctcaatgct cacgctgtag gtatctcagt tcggtgtagg     8100

tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct     8160

tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag     8220

cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga     8280

agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga     8340

agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg     8400

gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag     8460

aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag     8520

ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta aattaaaaat     8580

gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct     8640

taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac     8700

tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa     8760

tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac cagccagccg     8820

gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt     8880

gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca     8940
```

```
ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt    9000

cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg gttagctcct    9060

tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg    9120

cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactggtg    9180

agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg    9240

cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc atcattggaa    9300

aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt    9360

aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt    9420

gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt    9480

gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca    9540

tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat    9600

ttccccgaaa agtgccacct gacgtc                                        9626
```

```
<210>  47
<211>  22
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  47
gagaacccac tgcttactgg ct                                              22
```

```
<210>  48
<211>  21
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  48
gcattgctgt gcagtctgta a                                               21
```

```
<210>  49
<211>  7354
<212>  DNA
<213>  Artificial

<220>
<223>  recombinant DNA vector

<400>  49
cgcgattgct gatccccatg tgtatcactg gcaaactgtg atggacgaca ccgtcagtgc     60

gtccgtcgcg caggctctcg atgagctgat gctttgggcc gaggactgcc ccgaagtccg    120

gcacctcgtg cacgcggatt tcggctccaa caatgtcctg acggacaatg ccgcataac    180
```

agcggtcatt gactggagcg aggcgatgtt cggggattcc caatacgagg tcgccaacat   240

cttcttctgg aggccgtggt tggcttgtat ggagcagcag acgcgctact tcgagcggag   300

gcatccggag cttgcaggat cgccgcggct ccgggcgtat atgctccgca ttggtcttga   360

ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg cgcagggtcg   420

atgcgacgca atcgtccgat ccggagccgg gactgtcggg cgtacacaaa tcgcccgcag   480

aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg gaaaccgacg   540

ccccagcact cgtccgaggg caaaggaata gagtatgcag aaattgatga tctattaaac   600

aataaagatg tccactaaaa tggaagtttt tcctgtcata ctttgttaag aagggtgaga   660

acagagtacc tacattttga atggaaggat tggagctacg ggggtggggg tggggtggga   720

ttagataaat gcctgctctt tactgaaggc tctttactat tgctttatga taatgtttca   780

tagttggata tcataattta aacaagcaaa accaaattaa gggccagctc attcctccca   840

ctcatgatct atagatctat agatctctcg tgggatcatt gttttтctct tgattcccac   900

tttgtggttc taagtactgt ggtttccaaa tgtgtcagtt tcatagcctg aagaacgaga   960

tcagcagcct ctgttccaca tacacttcat tctcagtatt gttttgccaa gttctaattc   1020

catcagattt tagtggacat ctttattgtt taatagatca tcaatttctg catactaagc   1080

ttatcgatac cgtcgacctc gaggggggc ccggtaccca gcttttgttc cctttagtga   1140

gggttaattg cgcgcttggc gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat   1200

ccgctcacaa ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc   1260

taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga   1320

aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt   1380

attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg   1440

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc agggggataac   1500

gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg   1560

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca   1620

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc   1680

tccctcgtgc gctctcctgt ccgaccctg ccgcttaccg gatacctgtc gcctttctc   1740

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag   1800

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc   1860

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca   1920

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg   1980

aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg   2040

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct   2100

```
ggtagcggtg gttttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa    2160

gaagatcctt tgatctttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa    2220

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatccttttt aaattaaaaa    2280

tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc    2340

ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga    2400

ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca    2460

atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc    2520

ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat    2580

tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc    2640

attgctacag gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt    2700

tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc    2760

ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg    2820

gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt    2880

gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg    2940

gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga    3000

aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg    3060

taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg    3120

tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac acggaaatgt    3180

tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg ttattgtctc    3240

atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca    3300

tttccccgaa aagtgccacc taaattgtaa gcgttaatat tttgttaaaa ttcgcgttaa    3360

attttgtta aatcagctca ttttttaacc aataggccga atcggcaaa atcccttata    3420

aatcaaaaga atagaccgag atagggttga gtgttgttcc agtttggaac aagagtccac    3480

tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc    3540

cactacgtga accatcaccc taatcaagtt ttttggggtc gaggtgccgt aaagcactaa    3600

atcggaaccc taaagggagc ccccgattta gagcttgacg gggaaagccg cgaacgtgg    3660

cgagaaagga agggaagaaa gcgaaaggag cgggcgctag gcgctggca agtgtagcgg    3720

tcacgctgcg cgtaaccacc acacccgccg cgcttaatgc gccgctacag gcgcgtccc    3780

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat    3840

tacgccagct ggcgaagggg gatgtgctg caaggcgatt aagttgggta acgccagggt    3900

tttcccagtc acgacgttgt aaaacgacgg ccagtgagcg cgcgtaatac gactcactat    3960

agggcgaatt ggagctccac cgcggtgcgg ccgggccatg caggccagac atgataagat    4020
```

```
acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg    4080

aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa caagttgtcg    4140

agttacttgt acagctcgtc catgccgaga gtgatcccgg cggcggtcac gaactccagc    4200

aggaccatgt gatcgcgctt ctcgttgggg tctttgctca gggcggactg gtagctcagg    4260

tagtggttgt cgggcagcag cacggggccg tcgccgatgg gggtgttctg ctggtagtgg    4320

tcggcgagct gcacgctgcc gtcctcgatg ttgtggcgga tcttgaagtt caccttgatg    4380

ccgttcttct gcttgtcggc catgatatag acgttgtggc tgttgtagtt gtactccagc    4440

ttgtgcccca ggatgttgcc gtcctccttg aagtcgatgc ccttcagctc gatgcggttc    4500

accagggtgt cgccctcgaa cttcacctcg gcgcgggtct tgtagttgcc gtcgtccttg    4560

aagaagatgg tgcgctcctg gacgtagcct tcgggcatgg cggacttgaa gaagtcgtgc    4620

tgcttcatgt ggtcggggta gcgggcgaag cactgcaggc cgtagccgaa ggtggtcacg    4680

agggtgggcc agggcacggg cagcttgccg gtggtgcaga tgaacttcag ggtcagcttg    4740

ccgtaggtgg catcgccctc gccctcgccg gacacgctga acttgtggcc gtttacgtcg    4800

ccgtccagct cgaccaggat gggcaccacc ccggtgaaca gctcctcgcc cttgctcacc    4860

atggtggcgt tagcttgatt tgacagtggc tgggggttgc gccgccgggt tttataggaa    4920

gccacagcgg ccactcgagc cataaaaggc aactttagga acggcggggg gtgattggat    4980

tcgagtcgtt tattcaccgg ccttgccgca cagtgcagca ttttttacc ccctctcccc    5040

tccttttgcg ggggaaaaaa aaaaaaaaaa aaaaaaagg agaagagaaa aaagcgagc    5100

gagagagaaa gcgagattga ggaagaggat gaagagtttg gcgatgggtg ctggttccgt    5160

aggcccagat ggacaagaat agcccccgcc cttgcggaca gtatcccatt cagtgactca    5220

gatcagtcaa gcggccgctc tagacggcga tacagtcaac tgtctttgac cttttgttact    5280

actctcttcc gatgatgatg tcgcacttat tctatgctgt ctcaatgtta gaggcatatc    5340

agtctccact gaagccaatc tatctgtgac ggcatcttta ttcacattat cttgtacaaa    5400

taatcctgtt aacaatgctt ttatatcctg taaagaatcc attttcaaaa tcatgtcaag    5460

gtcttctcga ggaaaaatca gtagaaatag ctgttccagt ctttctagcc ttgattccac    5520

ttctgtcaga tgtgccctag tcagcggaga ccttttggtt ttgggagagt agcgacactc    5580

ccagttgttc ttcagacact tggcgcactt cggttttct ttggagcact tgagcttttt    5640

aagtcggcaa atatcgcatg cttgttcgat agaagacagt agcttcatgg tggcggatcc    5700

cgataagcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagacct    5760

cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg    5820

gaaagtcccg ttgatttgg tgccaaaaca aactcccatt gacgtcaatg gggtggagac    5880

ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt gatgtactgc caaaaccgca    5940
```

```
tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata    6000

aggtcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg    6060

gcgtacttgg catatgatac acttgatgta ctgccaagtg ggcagtttac cgtaaatact    6120

ccacccattg acgtcaatgg aaagtcccta ttggcgttac tatgggaaca tacgtcatta    6180

ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg cgggccattt accgtaagtt    6240

atgtaacgcg gaactccata tatgggctat gaactaatga ccccgtaatt gattactatt    6300

aataactagt caataatcaa tgtcaacgcg tatatctggc ccgtacatcg cgaagcagcg    6360

caaaacgcct aaccctaccc gggctgcagg aattctaccg ggtaggggag gcgcttttcc    6420

caaggcagtc tggagcatgc gctttagcag ccccgctggc acttggcgca tcacaagtgg    6480

cctctggcct cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggt    6540

ggccccttcg cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag    6600

ctcgcgtcgt gcaggacgtg acaaatggaa gtagcacgtc tcactagtct cgtgcagatg    6660

gacaagcacc gctgagcaat ggaagcgggt aggcctttgg ggcagcggcc aatagcagct    6720

tggctccttc gctttctggg ctcagaggct gggaaggggt gggtccgggg gcgggctcag    6780

gggcgggctc aggggcgggg cgggcgcgaa ggtcctccgg acccggcatt ctgcacgctt    6840

caaaagcgca cgtctgccgc gctgttctcc tcttcctcat ctccgggcct tcgacctgca    6900

tgaaaaagcc tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca    6960

gcgtctccga cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg    7020

taggagggcg tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc tacaaagatc    7080

gttatgttta tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg    7140

gggaattcag cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc    7200

aagacctgcc tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg    7260

cgatcgctgc ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa    7320

tcggtcaata cactacatgg cgtgatttca tatg                                 7354
```

<210> 50
<211> 10514
<212> DNA
<213> Artificial

<220>
<223> recombinant DNA vector

<400> 50
```
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc      60

attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga     120

gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc     180
```

```
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc    240

ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag    300

cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa    360

agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac    420

cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg    480

caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg    540

gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg    600

taaaacgacg gccagtgagc gcgcgtaata cgactcacta taggcgaat tggagctcca    660

ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca    720

aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc    780

tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt    840

ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct    900

tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca    960

gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc   1020

cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg   1080

ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc   1140

cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga   1200

acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct   1260

ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt   1320

agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg   1380

gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct   1440

cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga   1500

tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat   1560

ttgacagtgg ctggggggttg cgccgccggg ttttatagga agccacagcg gccactcgag   1620

ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg   1680

gccttgccgc acagtgcagc attttttttac cccctctccc ctccttttgc gggggaaaaa   1740

aaaaaaaaaa aaaaaaaaag gagaagagaa aaaagcgag cgagagagaa agcgagattg    1800

aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa    1860

tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg   1920

tctaacctag gagcccacca cagacactat actataacca cactgaattt ctctcctat    1980

ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc   2040

agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga   2100
```

```
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac    2160

aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc    2220

aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaagatca tgccattgca    2280

ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa    2340

acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt    2400

tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca    2460

caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg    2520

tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt    2580

agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt    2640

gcaaagtttc aatttttaat ctattgattt gaaagaatta ttatacatta ttttaaatat    2700

cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat    2760

taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt    2820

agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag    2880

agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa    2940

tctgtgggca catatttaaa ctctcaataa atattaactt tgattatta aaaaatacct    3000

ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa    3060

atgagaaaat tttcttctgc actgaattct tcttttaact aaatttaag aaaaaaacca    3120

atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag    3180

gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt    3240

accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt    3300

ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa    3360

attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc    3420

agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga    3480

atccataata aaatgtagaa gagtctgttg atctggggtg agttcagttt tctcctacac    3540

tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca    3600

ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg    3660

aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc    3720

acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat    3780

ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt    3840

cactgaaagt tcaagcttta attataacta atgaaaggtt ttcatatagt atgactacta    3900

tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt    3960

attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggattttat    4020
```

```
aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctacttttgc     4080

actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cctttaaaga     4140

aaaaaagttt gtcaacccct aaaatagaga gtcaacaatc tggacaattc agaccccaaa     4200

catctctttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg     4260

ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatataaa     4320

atactttttc ttaaataaat acaaaaatac gtatttgtaa aactaaatgt ggataaagtt     4380

ttctttgtca ataatgtaat aaatattgag aagggtgtgg tttggacaat catagcttac     4440

tggttttca gctaatggca aaataaaatt ccatcaaatg tcctgtgcat tatcacaatg     4500

tgtgtgtcta acaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt     4560

tcttgagaaa tttgtaaccc cagtgggtag gacttttctc tattcatcca catacgtgta     4620

ttcaccctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct     4680

tcccagtttg agactgtctt ctcaaaaga gggaaaaaat ggtcatcaat actcagttgg     4740

aaaaaatcca ttttatacac caaatcattt ctgctggcaa acactaggta gtaggactgg     4800

aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa     4860

tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca     4920

gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact     4980

cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt     5040

cgaggtcact tgtcgcaagt cacgaccttc actgtcttct ctagacggcg atacagtcaa     5100

ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg     5160

tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt     5220

attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatcct gtaaagaatc     5280

cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag     5340

tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggt     5400

tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgcact tcggtttttc     5460

tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag     5520

tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga     5580

gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg     5640

cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat     5700

tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat     5760

tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact     5820

gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac     5880

cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt     5940
```

```
gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta   6000

ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag   6060

gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg   6120

accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg   6180

cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc   6240

gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg   6300

cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc   6360

caaccggctc ccttctttgg tggccccttc gcgccacctt ctactcctcc cctagtcagg   6420

aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt   6480

ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg   6540

gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg   6600

tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg   6660

gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca   6720

tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt   6780

ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat   6840

ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg   6900

ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga   6960

ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc   7020

gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc   7080

cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg   7140

gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga   7200

ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg   7260

tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc   7320

tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg   7380

tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct   7440

tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc   7500

cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac   7560

tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg   7620

acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg   7680

cggccgtctg accgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca   7740

gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat aaacaataa   7800

agatgtccac taaaatggaa gttttttcctg tcatactttg ttaagaaggg tgagaacaga   7860
```

```
gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg tgggattaga    7920

taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt    7980

ggatatcata atttaaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat    8040

gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt    8100

ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc    8160

agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca    8220

gattttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc    8280

gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt    8340

aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct    8400

cacaattcca cacacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg    8460

agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct    8520

gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg    8580

gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc    8640

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg    8700

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct    8760

ggcgttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca    8820

gaggtggcga aacccgacag gactataaag ataccaggcg tttcccctg gaagctccct    8880

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc    8940

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt    9000

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc    9060

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc    9120

cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg    9180

gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc    9240

agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag    9300

cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga    9360

tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat    9420

tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag    9480

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat    9540

cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc    9600

cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat    9660

accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag    9720

ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg    9780
```

```
ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc      9840

tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca      9900

acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg      9960

tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc     10020

actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta     10080

ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc     10140

aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg     10200

ttcttcgggg cgaaaactct caaggatctt accgctgttg atccagtt cgatgtaacc      10260

cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc     10320

aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg cgacacgga aatgttgaat      10380

actcatactc ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag     10440

cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc     10500

ccgaaaagtg ccac                                                       10514
```

```
<210>   51
<211>   10247
<212>   DNA
<213>   Artificial

<220>
<223>   recombinant DNA molecule

<220>
<221>   misc_feature
<222>   (2626)..(2626)
<223>   unsequenced nucleotide


<400>   51
ctagaactag tggatccgga taagccagta agcagtgggt tctctagtta gccagagagc       60

tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgtca atggggcgga      120

gttgttacga cattttggaa agtcccgttg attttggtgc caaaacaaac tcccattgac      180

gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac gcccattgat      240

gtactgccaa aaccgcatca ccatggtaat agcgatgact aatacgtaga tgtactgcca      300

agtaggaaag tcccataagg tcatgtactg gcataatgc caggcgggcc atttaccgtc       360

attgacgtca ataggggcg tacttggcat atgatacact tgatgtactg ccaagtgggc       420

agtttaccgt aaatactcca cccattgacg tcaatggaaa gtccctattg gcgttactat      480

gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc agccaggcgg      540

gccatttacc gtaagttatg taacgcggaa ctccatatat gggctatgaa ctaatgaccc      600

cgtaattgat tactattaat aactagtcaa taatcaatgt caaccccggg ctgcaggaat      660
```

```
tctaccgggt aggggaggcg cttttcccaa ggcagtctgg agcatgcgct ttagcagccc      720

cgctggcact tggcgcatca caagtggcct ctggcctcgc acacattcca catccaccgg      780

tagcgccaac cggctccctt ctttggtggc cccttcgcgc caccttctac tcctccccta      840

gtcaggaagt tccccccgc cccgcagctc gcgtcgtgca ggacgtgaca aatggaagta      900

gcacgtctca ctagtctcgt gcagatggac aagcaccgct gagcaatgga agcgggtagg      960

cctttggggc agcggccaat agcagcttgg ctccttcgct ttctgggctc agaggctggg     1020

aaggggtggg tccgggggcg ggctcagggg cgggctcagg ggcggggcgg gcgcgaaggt     1080

cctccggacc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct gttctcctct     1140

tcctcatctc cgggccttcg acctgcatga aaaagcctga actcaccgcg acgtctgtcg     1200

agaagtttct gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggagggcg     1260

aagaatctcg tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata     1320

gctgcgccga tggtttctac aaagatcgtt atgtttatcg cactttgca tcggccgcgc     1380

tcccgattcc ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct     1440

cccgccgtgc acagggtgtc acgttgcaag acctgcctga aaccgaactg cccgctgttc     1500

tgcagccggt cgcggaggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg     1560

ggttcggccc attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat     1620

gcgcgattgc tgatccccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg     1680

cgtccgtcgc gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc     1740

ggcacctcgt gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa     1800

cagcggtcat tgactggagc gaggcgatgt cggggattc ccaatacgag gtcgccaaca     1860.

tcttcttctg gaggccgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga     1920

ggcatccgga gcttgcagga tcgccgcggc tccgggcgta tatgctccgc attggtcttg     1980

accaactcta tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc     2040

gatgcgacgc aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca     2100

gaagcgcggc cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac     2160

gccccagcac tcgtccgagg gcaaaggaat gcctgagaaa ggaagtgagc tgtaaaggct     2220

gagctctctc tctgacgtat gtagcctctg gttagcttcg tcactcactg ttcttgactc     2280

agcatggcaa tctgatgaaa tcccagctgt aagtctgcat aaattgatga tctattaaac     2340

aataaagatg tccactaaaa tggaagtttt ttactgtcat actttgttaa gaagggtgag     2400

aacagagtac ctacattttg aatggaagga ttggagctac ggggtgggg gtgggggtgg     2460

gattagataa atgcctgctc tttactgaag gctctttact attgctttat gataatgttt     2520

catagttgga tatcataatt taaacaagca aaccaaatt aagggccagc tcattcctcc     2580
```

```
actcatgatc tatagatcta tagatctctc gtgggatcat tgtttntctg atccactgga    2640

agcttatcga taccgtcgac ctcgagggggggcccggtac ccagcttttg ttccctttag    2700

tgagggttaa ttgcgcgctt ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt    2760

tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt    2820

gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg    2880

ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg    2940

cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg    3000

cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat    3060

aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc    3120

gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc    3180

tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt ccccctgga    3240

agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt    3300

ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg    3360

taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc    3420

gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg    3480

gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc    3540

ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg    3600

ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc    3660

gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct    3720

caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt    3780

taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa    3840

aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa    3900

tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc    3960

tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct    4020

gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca    4080

gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt    4140

aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt    4200

gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc    4260

ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc    4320

tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt    4380

atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact    4440

ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc    4500
```

```
ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt   4560

ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg   4620

atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct   4680

gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa   4740

tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt   4800

ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc   4860

acatttcccc gaaaagtgcc acctaaattg taagcgttaa tattttgtta aaattcgcgt   4920

taaatttttg ttaaatcagc tcattttta accataggc cgaaatcggc aaaatccctt    4980

ataaatcaaa agaatagacc gagatagggt tgagtgttgt tccagtttgg aacaagagtc   5040

cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg   5100

gcccactacg tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc cgtaaagcac   5160

taaatcggaa ccctaaaggg agcccccgat ttagagcttg acggggaaag ccggcgaacg   5220

tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag   5280

cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta cagggcgcgt   5340

cccattcgcc attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc   5400

tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag   5460

ggttttccca gtcacgacgt tgtaaaacga cggccagtga gcgcgcgtaa tacgactcac   5520

tatagggcga attggagctc caccgcggtg cggccgggcc atgcaggcca cgacatgata   5580

agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt   5640

tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt   5700

cttgtacagc tcgtccatgc cgagagtgat cccggcggcg gtcacgaact ccagcaggac   5760

catgtgatcg cgcttctcgt tggggtcttt gctcagggcg gactggtagc tcaggtagtg   5820

gttgtcgggc agcagcacgg ggccgtcgcc gatggggtg ttctgctggt agtggtcggc    5880

gagctgcacg ctgccgtcct cgatgttgtg gcggatcttg aagttcacct tgatgccgtt   5940

cttctgcttg tcggccatga tatagacgtt gtggctgttg tagttgtact ccagcttgtg   6000

ccccaggatg ttgccgtcct ccttgaagtc gatgcccttc agctcgatgc ggttcaccag   6060

ggtgtcgccc tcgaacttca cctcggcgcg ggtcttgtag ttgccgtcgt ccttgaagaa   6120

gatggtgcgc tcctggacgt agccttcggg catggcggac ttgaagaagt cgtgctgctt   6180

catgtggtcg gggtagcggg cgaagcactg caggccgtag ccgaaggtgg tcacgagggt   6240

gggccagggc acgggcagct tgccggtggt gcagatgaac ttcagggtca gcttgccgta   6300

ggtggcatcg ccctcgccct cgccggacac gctgaacttg tggccgttta cgtcgccgtc   6360

cagctcgacc aggatgggca ccaccccggt gaacagctcc tcgcccttgc tcaccatggt   6420
```

```
ggcgttagct tgatttgaca gtggctgggg gttgcgccgc cgggttttat aggaagccac    6480

agcggccact cgagccataa aaggcaactt taggaacggc ggggggtgat tggattcgag    6540

tcgtttattc accggccttg ccgcacagtg cagcattttt ttaccccctc tcccctcctt    6600

ttgcggggga aaaaaaaaa aaaaaaaaa aaaggagaag agaaaaaaag cgagcgagag    6660

agaaagcgag attgaggaag aggatgaaga gtttggcgat gggtgctggt tccgtaggcc    6720

cagatggaca agaatagccc ccgcccttgc ggacagtatc ccattcagtg actcagatca    6780

gatcaagcgg ccgccagtgt gatggatatc tgcagaattc cagcacactg gcggccgtta    6840

ctagtggatc ctcaatcagt ccatcacctg ggtacttttt ctacccaatg ctgaaagtga    6900

atatttgaga aactatgttt ctaactggaa gaattctgaa actatgttct tctacaacat    6960

gtagacgatg tgtcacaatg aaaatgaaag ttccttctga agagctgtaa ccctacatag    7020

cctaggatgc catttatact ttctctagcc cagaggctct gtatttctgc tctcaggcta    7080

cattactaca taaccatcaa accaactgcc cgaaattgct gagcaattga tcataacata    7140

cagacctgat tcagatatat aaattacatt tttattttta tcgtaccaaa acatacaggt    7200

gaaaactttc aaattcaggc aactgacatg ccaggtgcct acctatctgt gaagaccaga    7260

gcactgtaag acacggaata gctggttgaa aagtgaaaaa tccagtgagt taggttgttt    7320

gattatacag aatcatagaa ttttcaaagt gaaaggacat ggtgatcact gatttcagtg    7380

actatccaac aacattagca agagcccta gggttcccta ccagttcctt cagaagcctc    7440

cttcaaagct gtgagtgcat ggccaaaggg gaagctacat tcatcaacca gactaatcag    7500

agaccaggga gcagaaataa ggttaacacg gaatggagta catttgcttt gggggacgtt    7560

catgttttct tgcgcatact gaaaacagct gggtgctgat tttggtataa cactcagagc    7620

ctcttttgat aaaaaggcat tctatatgaa tgaaaaccct gttggcttca cattggtcaa    7680

agcaatggat tttaaatagt gactcaactt cctagaaaac tttgtattta gaaatgtttt    7740

cttcatacag ctaacaggca aaactaatca aaaccttgtt gttagtctga tgtttatttt    7800

ctgtgcagca aacaaactcc cataaaaacc tctgaggttt tcttagttat gtttgcaatt    7860

ttttaagtgt tgaactgcct aagtctctga tgtcatgaca cttaattcat atattttctt    7920

gcttttgcta gaactgaaca aacactccct atcaagattg accctctcta gaaataaaat    7980

aaattgtcat atgaagtcaa aatatgacct atcaagttta tgatagcttt ctttgacatt    8040

tatgggagga gaaaattata ccatccatac tacataggtt ggggtttaaa taatatttgc    8100

ctataaacaa taagcaattg cagattgatt ttaagtaaac gtatgcaatg cctgtaagca    8160

acaggaaata tattttacat caatgtatgc actgcctgta agtaacagga attatatttt    8220

acatgaatgt gtgaagtgat aatttttttaa taggcaaaaa ttagtgttaa aagcggaact    8280

ggcatgcttt cctaggctgt ctcagtgata atgacctcaa agtcaatgat ttcttcatga    8340
```

```
agatgaaata agtttattgc aacttgattg catcacctaa aatacaaccc ttgacttgga    8400

agcttgccct tatgttttat gaggtctgaa ataaaaagga aactgatgga actgaaccag    8460

caaattattt tgcaatccat gactacagtt gtatacagca gatgtagtca attgttgagg    8520

aaagttggac caaatttcca tcctgttcag tactgctgat agctctattc aaaatacatt    8580

cttgccttta gaggaactga cccattagta atcatacaga tagaattttg tgaactattt    8640

ttaaagggta ggttattaat caatgagtga ctctcaaata tcttgatggt ttccactgtg    8700

aaaaccaagt atctgcatga tagagcaaaa atttaatttt ctcagatgct ttagagtaca    8760

tatagatgaa gactgccttc ttactagatc agttttgaga gagatgaaaa caaaagaaag    8820

aaaaaatagg atagctgatg aaatgaaact tcaaatttaa tttgatcatg ttggcatttt    8880

ttattccttg agtgcattga aaataactga ttaatactag acacatataa tactatcaag    8940

ttacacctaa aatttatttt gtcagccagg caaccaatca gaaccaaata gtgaccaaga    9000

attatttttt caaacacgtt tgtcttccag gaccttactg agttagttcc acagatcatg    9060

aaaacatggg tcttttctca ggtcttctac atcatttccc ttaagttgac acccaaagtc    9120

ataagcttga ctttagacac aggtttcaaa agtagaaagt tttattctct tatgtttcta    9180

tggctttgaa gttttgcaga actataaaaa taatttgttt tacccttttcc acaactcatg    9240

gacgttcctt gaagtgaagg aaacgaatac tcgtctttgc ataataaaca gcaaaaaaaa    9300

aaaaaaaaaa aaaattaaat gaatctaaat tctcatatgc catttaaagt ttcagaccca    9360

gacatcttca ggattttctt acatttaaag cagaatgaca ctactgatac acaaaagtaa    9420

attttaattc tgttttttaat gcatgaggca atataaaaat taatcatgat gatcattata    9480

tttgttctaa cactaaagtt tttcccataa tcccccagat ttttcaagtc tcttttctcc    9540

tcatggcaga tgcccaacca cagagagaaa ggttctataa taaaacttta aaaacactga    9600

cctaaaagag aaatgtttta tcgccatgca atattctaaa ttctggcatt tggtggaaaa    9660

actgtacaca ttttttggcaa tggctttagt gtccactggt ctggcagcta taatgagaat    9720

aataaagtat atggtgatga ctgtttactt accattatga gacatcccca ctgcggtacc    9780

cggcgataca gtcaactgtc tttgaccttt gttactactc tcttccgatg atgatgtcgc    9840

acttattcta tgctgtctca atgttagagg catatcagtc tccactgaag ccaatctatc    9900

tgtgacggca tctttattca cattatcttg tacaaataat cctgttaaca atgctttat    9960

atcctgtaaa gaatccattt tcaaaatcat gtcaaggtct tctcgaggaa aaatcagtag   10020

aaatagctgt tccagtcttt ctagccttga ttccacttct gtcagatgtg ccctagtcag   10080

cggagacctt ttggttttgg gagagtagcg acactcccag ttgttcttca gacacttggc   10140

gcacttcggt ttttctttgg agcacttgag cttttttaagt cggcaaatat cgcatgcttg   10200

ttcgatagaa gacagtagct tcatggtggc aagcttaagt ttaaacg                 10247
```

```
<210>   52
<211>   10514
<212>   DNA
<213>   Artificial

<220>
<223>   recombinant DNA vector

<400>   52
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc      60

attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga     120

gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc     180

caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc     240

ctaatcaagt tttttggggt cgaygtgccg taaagcacta aatcggaacc ctaaagggag     300

cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa     360

agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac     420

cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg     480

caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg     540

gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg     600

taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca     660

ccgcggtgcg ccgggccat gcaggccaga catgataaga tacattgatg agtttggaca     720

aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc     780

tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt     840

ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct     900

tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca     960

gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc    1020

cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg    1080

ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc    1140

cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga    1200

acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct    1260

ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt    1320

agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg    1380

gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct    1440

cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga    1500

tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat    1560

ttgacagtgg ctggggggttg cgccgccggg ttttatagga agccacagcg gccactcgag    1620
```

```
ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg    1680

gccttgccgc acagtgcagc attttttac cccctctccc ctccttttgc gggggaaaaa    1740

aaaaaaaaaa aaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg    1800

aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa    1860

tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg    1920

tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat    1980

ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc    2040

agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga    2100

ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac    2160

aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc    2220

aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaagatca tgccattgca    2280

ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa    2340

acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt    2400

tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca    2460

caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg    2520

tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt    2580

agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt    2640

gcaaagtttc aatttttaat ctattgattt gaaagaatta ttatacatta ttttaaatat    2700

cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat    2760

taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt    2820

agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag    2880

agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa    2940

tctgtgggca catatttaaa ctctcaataa atattaactt tgattattta aaaaatacct    3000

ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa    3060

atgagaaaat tttcttctgc actgaattct tcttttaact aaattttaag aaaaaaacca    3120

atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag    3180

gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt    3240

accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt    3300

ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa    3360

attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc    3420

agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga    3480

atccataata aaatgtagaa gagtctgttg atctggggtg agttcagttt tctcctacac    3540
```

```
tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca    3600

ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg    3660

aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc    3720

acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat    3780

ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt    3840

cactgaaagt tcaagctttta attataacta atgaaaggtt ttcatatagt atgactacta    3900

tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt    3960

attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggattttat    4020

aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctactttttgc    4080

actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cctttaaaga    4140

aaaaaagttt gtcaacccct aaaatagaga gtcaacaatc tggacaattc agaccccaaa    4200

catctctttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg    4260

ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatataaa    4320

atactttttc ttaaataaat acaaaaatac gtatttgtaa aactaaatgt ggataaagtt    4380

ttctttgtca ataatgtaat aaatattgag aagggtgtgg tttggacaat catagcttac    4440

tggttttca gctaatggca aaataaaatt ccatcaaatg tcctgtgcat tatcacaatg    4500

tgtgtgtcta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt    4560

tcttgagaaa tttgtaaccc cagtgggtag gacttttctc tattcatcca catacgtgta    4620

ttcaccctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct    4680

tcccagtttg agactgtctt ctcaaaaaga gggaaaaaat ggtcatcaat actcagttgg    4740

aaaaaatcca ttttatacac caaatcattt ctgctggcaa acactaggta gtaggactgg    4800

aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa    4860

tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca    4920

gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact    4980

cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt    5040

cgaggtcact tgtcgcaagt cacgaccttc actgtcttct ctagacggcg atacagtcaa    5100

ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg    5160

tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt    5220

attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatcct gtaaagaatc    5280

cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag    5340

tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggt    5400

tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgcact tcggttttttc    5460
```

```
tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag   5520

tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga   5580

gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg   5640

cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat   5700

tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat   5760

tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact   5820

gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac   5880

cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt   5940

gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta   6000

ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag   6060

gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg   6120

accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg   6180

cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc   6240

gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg   6300

cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc   6360

caaccggctc ccttctttgg tggccccttc gcgccacctt ctactcctcc cctagtcagg   6420

aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt   6480

ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg   6540

gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg   6600

tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg   6660

gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca   6720

tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt   6780

ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat   6840

ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg   6900

ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga   6960

ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc   7020

gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc   7080

cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg   7140

gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga   7200

ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg   7260

tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc   7320

tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg   7380
```

```
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct    7440

tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc    7500

cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac    7560

tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg    7620

acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg    7680

cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca    7740

gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa    7800

agatgtccac taaaatggaa gtttttcctg tcatactttg ttaagaaggg tgagaacaga    7860

gtacctacat tttgaatgga aggattggag ctacgggggt ggggtggggg tgggattaga    7920

taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt    7980

ggatatcata atttaaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat    8040

gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt    8100

ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc    8160

agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca    8220

gattttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc    8280

gataccgtcg acctcgaggg ggggcccggt acccagcttt gttccctttt agtgagggtt    8340

aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct    8400

cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg    8460

agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct    8520

gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg    8580

gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc    8640

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg    8700

aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct    8760

ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca    8820

gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct    8880

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc    8940

gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt    9000

tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc    9060

cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc    9120

cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg    9180

gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc    9240

agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag    9300
```

```
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga    9360

tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat    9420

tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag    9480

ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat    9540

cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc    9600

cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat    9660

accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag    9720

ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg    9780

ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc    9840

tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca    9900

acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg    9960

tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc   10020

actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta   10080

ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc   10140

aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg   10200

ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc   10260

cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc   10320

aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat   10380

actcatactc ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag   10440

cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc   10500

ccgaaaagtg ccac                                                     10514


<210>   53
<211>   168
<212>   DNA
<213>   Homo sapiens

<400>   53
cggaggacag tactccgctc ggaggacagt actccgctcg gaggacagta ctccgctcga    60

ggacagtact ccgctcggag acagtactc cgatccgtcg actctagtcc gctcggagga   120

cagtactccg ctcggaggac agtactccgc tcggaggaca gtactccg              168


<210>   54
<211>   22
<212>   DNA
<213>   Artificial

<220>
<223>   primer
```

<400> 54
gagaacccac tgcttactgg ct
<div style="text-align:right">22</div>

<210> 55
<211> 20
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 55
caagatcgcc ctcgcctttg
<div style="text-align:right">20</div>

<210> 56
<211> 33
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 56
atatggtacc gcagtgggga tgtctcataa tgg
<div style="text-align:right">33</div>

<210> 57
<211> 30
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 57
aattggatcc tcaatcagtc catcacctgg
<div style="text-align:right">30</div>

<210> 58
<211> 30
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 58
cgtctagaga gtgtgtcctg tcagaagaac
<div style="text-align:right">30</div>

<210> 59
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer

<400> 59
atgcggccgc actcctgacc tcaggtgatc c
<div style="text-align:right">31</div>

```
<210>  60
<211>  29
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  60
cgtctagaga agacagtgaa ggtcgtgac                                29


<210>  61
<211>  29
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  61
gcgcggccgc gtctaaccta ggagcccac                                29    .


<210>  62
<211>  28
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  62
gctctagagg acgaatgcca agatctga                                 28


<210>  63
<211>  29
<212>  DNA
<213>  Artificial

<220>
<223>  primer

<400>  63
caagcggccg ccagtgtgat ggatatctg                                29


<210>  64
<211>  5218
<212>  DNA
<213>  Artificial

<220>
<223>  recomkbinant DNA

<400>  64
gccggagcgg tcgagttctg gaccgaccgg ctcgggttct cccgggactt cgtggaggac    60

gacttcgccg gtgtggtccg ggacgacgtg accctgttca tcagcgcggt ccaggaccag    120

gtggtgccgg acaacaccct ggcctgggtg tgggtgcgcg gcctggacga gctgtacgcc    180
```

```
gagtggtcgg aggtcgtgtc cacgaacttc cgggacgcct ccgggccggc catgaccgag      240

atcggcgagc agccgtgggg gcgggagttc gccctgcgcg acccggccgg caactgcgtg      300

cacttcgtgg ccgaggagca ggactgacac gtgctacgag atttcgattc caccgccgcc      360

ttctatgaaa ggttgggctt cggaatcgtt ttccgggacg ccggctggat gatcctccag      420

cgcggggatc tcatgctgga gttcttcgcc caccccaact tgtttattgc agcttataat      480

ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttcactgcat      540

tctagttgtg gtttgtccaa actcatcaat gtatcttatc atgtctgtat accgtcgacc      600

tctagctaga gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg      660

ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa      720

tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac      780

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt      840    .

gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg ctgcggcga      900

gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca      960

ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg     1020

ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt     1080

cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc     1140

ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct     1200

tcgggaagcg tggcgctttc tcaatgctca cgctgtaggt atctcagttc ggtgtaggtc     1260

gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta     1320

tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca     1380

gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag     1440

tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag     1500

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt     1560

agcggtggtt ttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa     1620

gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg     1680

attttggtca tgagattatc aaaaaggatc ttcacctaga tcctttttaaa ttaaaaatga     1740

agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta     1800

atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc     1860

cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg     1920

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga     1980

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt     2040

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt     2100
```

```
gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc   2160

caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc   2220

ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca   2280

gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag   2340

tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg   2400

tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa   2460

cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa   2520

cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga   2580

gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga   2640

atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg   2700

agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt   2760

ccccgaaaag tgccacctga cgtcgacgga tcgggagatc cgatgttttc tgagttactt   2820

ttgtatcccc accccccctc gaggatcgat ccgctcggag gacagtactc cgctcggagg   2880

acagtactcc gctcggagga cagtactccg ctcgaggaca gtactccgct cggaggacag   2940

tactccgatc cgtcgactct agtccgctcg gaggacagta ctccgctcgg aggacagtac   3000

tccgctcgga ggacagtact ccgctcgagg acagtactcg attcgaggac attttgacac   3060

ccccataata tttttccaga attaacagta taaattgcat ctcttgttca agagttccct   3120

atcactctct ttaatcacta ctcacagtaa cctcaactcc tgcccaagct tggtaccacc   3180

atggacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg   3240

ggttacatcg aactggatct caacagcggt aagatccttg agagttttcg ccccgaagaa   3300

cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt   3360

gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag   3420

tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt   3480

gctgccataa ccatgagtga taacactgcg gccaacttac ttctgacaac gatcggagga   3540

ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt   3600

tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta   3660

gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg   3720

caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc   3780

cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt   3840

atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat ctacacgacg   3900

gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg   3960

attaagcatt ggtaatctag agagggcccg tttaaacccg ctgatcagcc tcgactgtgc   4020
```

```
cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg accctggaag   4080

gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta   4140

ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag   4200

acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg gaaagaacca   4260

gctggggctc taggggtat ccccacgcgc cctgtagcgg cgcattaagc gcggcgggtg   4320

tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg   4380

ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg   4440

gcatcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt   4500

agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt   4560

tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta   4620

tctcggtcta ttcttttgat ttataaggga ttttggggat ttcggcctat tggttaaaaa   4680

atgagctgat ttaacaaaaa tttaacgcga attaattctg tggaatgtgt gtcagttagg   4740

gtgtggaaag tccccaggct ccccaggcag gcagaagtat gcaaagcatg catctcaatt   4800

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca   4860

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa   4920

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag   4980

aggccgaggc cgcctctgcc tctgagctat tccagaagta gtgaggaggc ttttttggag   5040

gcctaggctt ttgcaaaaag ctcccgggag cttgtatatc cattttcgga tctgatcagc   5100

acgtgttgac aattaatcat cggcatagta tatcggcata gtataatacg acaaggtgag   5160

gaactaaacc atggccaagt tgaccagtgc cgttccggtg ctcaccgcgc gcgacgtc    5218
```